2021

NON-MARKET VALUATION OF UTILITY-SCALE SOLAR ARRAYS

Vasundhara Gaur
University of Rhode Island, vasundharagaur@gmail.com

NON-MARKET VALUATION OF UTILITY-SCALE

SOLAR ARRAYS

BY

VASUNDHARA GAUR

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

ENVIRONMENTAL AND NATURAL RESOURCE ECONOMICS

UNIVERSITY OF RHODE ISLAND

2021

DOCTOR OF PHILOSOPHY DISSERTATION

OF

VASUNDHARA GAUR

APPROVED:

Dissertation Committee:

Major Professor Corey Lang

Todd Guilfoos

David Bidwell

Brenton DeBoef
DEAN OF GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND
2021

ABSTRACT

In this dissertation, I investigate preferences for solar energy and solar siting decisions in the states of Rhode Island (RI) and Massachusetts (MA). There is broad support for solar energy in the United States, but the construction of utility-scale (1 megawatt capacity and above) installations has become controversial in recent years because of the potential local land use externalities that can result. Despite the vast literature on the impacts of various energy-producing infrastructures on areas in their vicinity, surprisingly little research exists on the externalities associated with utility-scale solar arrays within the United States. Estimating the value that people place on solar development will be valuable for informing policy.

In Manuscript 1, I quantify the externalities from nearby solar arrays in MA and RI using the hedonic method and a difference-in-differences, repeat sales identification strategy. I find that property prices for homes lying within 0.6 miles of a solar installation decline between 1.5% and 3.6% post array construction. Results also suggest that this effect is driven by solar developments on farm and forested lands and in rural areas, which is intuitive given the composite impact of loss of open space and loss of rural character. For these states, local disamenities are of the same order of magnitude as the global benefits of abated carbon emissions, which helps explain local opposition to siting.

In Manuscript 2, I complement the revealed preference hedonic analysis of Manuscript 1 with a stated preference analysis to provide further insight into the preferences for the following solar siting attributes: size of installation, visibility, setback distance, probability of future residential development, and current land use. Using data gathered from a survey of 656 respondents in RI, I find that land use is the primary determinant of public approval of solar development. Respondents are willing to pay an additional $10 to $21 in monthly electricity bills for solar development on commercial and brownfield sites, and between $13 and $49 to avoid developments on farm and forested land. Additionally, they prefer installations that are completely hidden from view and are willing to pay between $6 and $8 per month for a solar array to be fully hidden behind a vegetative screen.

In the third and final manuscript, I examine the differences in preferences and willingness to pay (WTP) magnitudes for solar siting attributes in RI between a random sample of the population and a convenience sample of engaged stakeholders. Engaged respondents are recruited from a list of individuals who had registered for a webinar titled "Valuing Siting Options for Solar Energy in RI" that was organized by the University of Rhode Island and advertised on social media in August 2021. The random survey sample used for analysis in Manuscript 2 is comprised of 656 respondents. I find that the preferences of both the engaged and random sample respondents are similar for most attributes. However, there are large differences in WTP magnitudes, with engaged respondents exhibiting WTP values that are two to four times higher than those of the random sample respondents. Although the overall preferences can be said to be representative of the population at large, caution should be exercised when generalizing valuation estimates derived from convenience samples.

ACKNOWLEDGEMENTS

"Recalling the little amount of water it was given in its early age, the coconut tree bears the nectar [coconut water] on its head for the rest of its life. A wise man [similarly] should never forget the help he has received."

- Sanskrit epigram

This dissertation is the culmination of five years of work and research, which would not have been possible without support from several people. I could conceivably write a chapter-length acknowledgements section thanking everyone for all the things that they need to be thanked for, but I will attempt to be brief.

First and foremost, I would like to thank my advisor Corey Lang for his teachings, guidance, and patience with everything (but especially with my less than stellar writing skills). You have been very generous with your time, swiftly responding to all emails and being available to meet and/or chat at the shortest notice. You are always thinking ahead, and I would not be half as good at meeting deadlines if it was not for your forethought and planning. You took time out of your busy schedule to help me implement a survey in the middle of a pandemic, assisting me with stuffing and stamping envelopes when you could have been preparing for classes. In a year when the world shut down and research output was at an all-time low, you made sure that I had a job market paper ready on time, and I am eternally grateful for that. Observing how you think about and do research has been an amazing learning experience for me that I hope to emulate in the future.

I would also like to extend my gratitude to other faculty and staff members at URI – to my committee members Todd Guilfoos and David Bidwell, for their constructive feedback; to Simona Trandafir, for funding my initial years as a graduate student and for numerous helpful conversations and guiding advice ever since; to Emi and Hiro Uchida for your teachings and the delightful food and company during Thanksgiving each year; to Jim Opaluch and Tom Sproul, for your instruction in Microeconomics; to Kate Venturini, for helping with research extension; and lastly, to Ben Morris and Deb Bourassa, for their administrative assistance and for making sure I submitted my timesheets on time.

Two out of three chapters of this dissertation are based on a survey, which would not have been possible without the assistance of several people. To my co-author Greg Howard, I am grateful for your help with survey design, choice analysis (you basically taught me how to do a Choice Experiment), and manuscript writing. Next, Ruth Quainoo, my other co-author, thank you for taking over the mind-numbing task of data entry. I'm also grateful to Carolyn Lang, Sadie Lang, and Vivian Lang for spending what could have been precious family time helping me arrange, stuff, seal, and stamp envelopes. I'm thankful for undergraduate Energy Fellow Molly Perkins for assisting with the first round of envelope stuffing and for my roommate Jodi Sutherland during the final round of envelope stuffing. And finally, to my fellow graduate students Luran Dong, Natalie Meyer, and Jarron VanCeylon, thank you for taking time out of your weekend to help me stamp and seal envelopes.

Words fall short in expressing my utmost gratitude towards all past and current graduate students at ENRE. It has been an amazing five years, and a lot of that can be attributed to the company I had in Coastal Institute Room 222. To my previous roommate, Priya Thomas, thank you for keeping me sane during the Covid lockdown, and for supporting me both at the office and at home, beginning with your helpful emails from before I came to URI, right to this day (one year after you have already graduated). Special mention to the people in "our corner" – Mike Weir, Ben Blachly, Jason Walsh, and Sarah Hayden. Each conversation and experience was twice as epic because it was shared with you guys, and exams and comps were only half as daunting because of your constant presence and support.

I also thank three anonymous referees, Ben Hoen, Salma Elmallah, and conference participants at AERE, NAREA, and USDA working group W4133 for their constructive feedback. This work was supported by the USDA National Institute of Food and Agriculture, Agricultural and Food Research Initiative Competitive Program, Critical Agricultural Research and Extension, grant number 2019-68008-29826.

Last, but by no means the least, I would like to thank my family – my parents, brother, and sister-in-law, for their love and affection which gave me strength to travel alone and pursue a PhD far away from the comfort of home. Your moral support allowed me to persevere and continue working hard through the pandemic.

PREFACE

This dissertation is written in three-manuscript form. The first manuscript is co-authored with Corey Lang and is under review by the *Journal of Environmental Economics and Management*. The second manuscript is co-authored with Corey Lang, Gregory Howard, and Ruth Quainoo. It is being reviewed by the *Resource and Energy Economics* journal. The third manuscript is also co-authored with Corey Lang, Gregory Howard, and Ruth Quainoo. It is being prepared for submission to the *Applied Economics Letters* journal.

Manuscript 1: House of the Rising Sun: The Effect of Utility-scale Solar Arrays on Housing Prices

Manuscript 2: When Energy Issues are Land Use Issues: Estimating Preferences for Utility-Scale Solar Energy Siting

Manuscript 3: Are the Loudest Voices in the Room Different? Comparing Engaged Versus Random Samples in a Contingent Valuation Framework

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Manuscript – 1

Manuscript – 2

Manuscript – 3

Manuscript – 1

Submitted to the Journal of Environmental Economics and Management, August 2020

House of the Rising Sun:
The Effect of Utility-scale Solar Arrays on Housing Prices

Vasundhara Gaur and Corey Lang
Department of Environmental and Natural Resource Economics, University of Rhode Island

ABSTRACT

While utility-scale solar energy is important for reducing dependence on fossil fuels, solar arrays use significant amounts of land (about 5 acres per MW of capacity) and may create local land use disamenities. This paper seeks to quantify the externalities from nearby solar arrays using the hedonic method. We study the states of Massachusetts and Rhode Island, which have high population densities and ambitious renewable energy goals. Using difference-in-differences, repeat sales identification strategies, results suggest that houses within 0.6 miles depreciate 1.5-3.6% following construction of a solar array. However, additional analysis reveals that this average effect is primarily driven by solar developments on farm and forest lands and in rural areas, which is intuitive given the composite impact of solar, loss of open space, and loss of rural character. For these states, the local disamenities are the same order of magnitude as the global benefits of abated carbon emissions, which helps explain local opposition to siting.

Keywords: Solar energy; Utility-scale solar; Hedonic valuation; Difference-in-differences

JEL codes: Q24; Q42; Q51

1 INTRODUCTION

Solar energy in the United States has grown at an average rate of 49% per year since 2009, making the US the second largest producer of solar energy in the world (EIA International Energy Outlook 2019). In 2019, solar energy accounted for 40% of all new capacity additions in the country, the largest ever in its history, and exceeding all other energy sources (Perea et al., 2020). By June 2020, the cumulative installed capacity of solar in the United States reached 81.4 gigawatts (GW), which is enough to power 15.7 million homes (Perea et al., 2020). Solar is predicted to overtake wind to become the largest source of renewable energy in the US by 2050, accounting for 46% of all energy produced from renewable sources (EIA Annual Energy Outlook 2018).

While there is a broad support for renewable energy in the United States (Bates and Firestone, 2015; Farhar, 1994; Firestone et al., 2018; Hoen et al., 2019; Krohn and Damborg, 1999), and for solar energy in particular (Carlisle et al., 2014, 2015; Farhar, 1994; Greenberg, 2009; Jacobe, 2013; Pew Research Center, 2019), the development of large-scale solar installations has not been obstacle free. One major hurdle to overcome before construction begins is the siting process. Solar installations require over ten times more land area than non-renewable sources to generate the same amount of energy, and the requirement of large tracts of land for their construction has become the largest cause of land use change in the United States (Trainor et al. 2016; Ong et al. 2013). Recently, the siting of large solar projects has become contentious in some parts of the country due to concerns about visual disamenities, impacts on ecosystems, building new transmission lines, loss of a town's rural character, water pollution, fire risk, water use, and reduction in property values (Farhar et al., 2010; Gross, 2020; Lovich and Ennen, 2011). The debate is especially heated when solar development is proposed on existing farm and forest lands, which is common because these are the cheapest locations for development, but many consider antithetical to environmental objectives (Kuffner, 2018; Naylor, 2019).

The purpose of this paper is to quantify the externalities associated with proximity to utility-scale solar installations using hedonic valuation. Theory indicates that property values will reflect people's willingness to pay (WTP) to avoid the cumulative disamenities of solar development (Bishop et al., 2019; Rosen, 1974). Our objective is to provide policy relevant non-market cost estimates in order to help state and municipal policy makers implement policies and decisions that reflect public preferences.

We focus on the states of Massachusetts (MA) and Rhode Island (RI), which are ideal for two reasons. First, both states have recently experienced a sudden boom in the development of large-scale solar installations. This trend has been driven by the Renewable Portfolio Standards (RPS), regulations that require increased energy production from renewable energy sources, which have been adopted by both states. MA's RPS calls for 25% of electricity generated by renewable sources by 2030 and RI's RPS calls for 38.5% by 2035. Second, both states have high population density, ranked 2nd and 3rd among U.S. states. This level of development means that most solar sites are proximate to residential areas, which yields many observed transactions for precise estimates.

We analyze the impact of utility-scale solar installations sized 1 MW and above on nearby property prices in MA and RI.[1] We apply two empirical approaches. First, we use a traditional repeat sales, difference-in-differences (DID) identification strategy, which compares changes in housing prices after construction for nearby properties with those further away. We empirically estimate the spatial extent of treatment to be 0.6 miles from the solar installation and choose a cutoff for control properties of two miles. Our primary sample consists of 282 solar installations, 11,292 housing transactions occurring within 0.6 miles (treated group), and 95,999 transactions between 0.6 and two miles (never-treated control group). However, pre-treatment trends are not perfectly parallel and there exist differences in the housing stock between treatment and control, which raise concerns about necessary assumptions holding. Given these concerns, we also estimate a DID model using only ever-treated properties, which relies entirely on temporal variation in construction dates. This method is preferred if there are endogeneity concerns about the siting of solar being correlated with trends in prices and not just levels. We present both models

[1] Following the U.S. Energy Information Administration (EIA), we define large-scale solar installations as those with an installed capacity of 1 MW or larger.

for all specifications and hedge about which is preferred.

Across a variety of specifications, our results suggest that solar installations negatively affect nearby property values. Results that average effects across all sites find negative impacts ranging from -1.5% to -3.6%, with the models using only the ever-treated sample consistently indicating larger effects. However, we examine heterogeneity in treatment effects that lead to important insights. We posit that solar arrays on farm and forest lands ("greenfields") cause greater externalities, given the combination of solar-specific disamenities and loss of open space amenities. Further, rural areas may be more impacted by solar if industrial solar arrays are incongruent with highly valued rural character, but on the other hand space is scarcer in non-rural areas. We find that the average treatment effects are to a large part driven by greenfield arrays and arrays in rural areas. Coefficients on non-greenfield sties and non-rural sites are consistently negative, but never statistically significantly different than zero.

Our findings suggest that utility-scale solar arrays create local, negative externalities. This helps explain local concerns and opposition to new development and gives pause to current practices of not including proximate residents in siting decisions or compensating them after siting has occurred. While a full benefit-cost analysis is well beyond the scope of this paper, we can compare the local, negative externalities to the value of greenhouse gas reductions from the solar arrays, which is the major global benefit. Our back-of-the-envelope calculations imply a benefit-cost ratio ranging from 1.65 to 0.69 depending on the choice of model. While it is promising that the benefit-cost ratio can be greater than one, it is clear regardless that a substantial and uneven burden is imposed on local areas to achieve global benefits of a similar magnitude. However, benefit-cost ratios are likely to be more favorable in other states due to different sources of fossil fuels and sparser population.

The recent growth in utility-scale solar has been met with a wave of research focused on assessing externalities and siting preferences. Prior hedonic valuation research includes Abashidze (2019) who applies a DID methodology with treatment and control defined by proximity, similar to our first model. Using data from North Carolina, USA, she finds that property values decline 8.7% post-construction within 1 street-network mile of a solar array. She similarly tests for treatment effect heterogeneity by prior land use, but finds no statistical differences. Dröes and Koster (2021) apply DID and rely on an ever-treated only sample, similar to our second model. Working in the context of the Netherlands, they find a 2-3% decline in value for properties within 1 km of a solar array. Jarvis (2021) uses data from the United Kingdom and also applies a DID methodology, but uses properties near solar sites that were proposed but not built as the control group. He finds zero statistical impact on property values.[2]

In addition to the hedonic valuation studies, there are several stated preference studies that also examine externalities from utility-scale solar siting. Botelho et al. (2017) survey residents of Portugal using a contingent valuation approach and find that respondents are willing to accept \$12.93 – \$56.64 per month on average as compensation for being in the vicinity of large solar installations, which is on par with our WTP estimates. In addition, Botelho et al. conduct a discrete choice experiment to delve into aspects of siting that drive the disamenity and estimate that respondents are willing to pay \$8.65, \$7.57, and \$5.15 per month to avoid negative impacts on flora and fauna, landscape, and glare effects, respectively. Kim et al. (2020) carry out a choice experiment in South Korea focused on land use and find large WTP (\$1,000-\$2,000 per household per month) for solar to be sited on rooftops and walls instead of farmlands, orchards, and mountainous areas. Lang et al. (2021) develop a choice experiment focused also on land use, but also attributes of arrays such as visibility and

---

[2] Our work is additionally closely related to the extensive hedonic applications assessing externalities of wind energy. Within the United States, studies that use data with large numbers of observations close to turbines find no significant impact on property prices, including Hoen and Atkinson-Palombo (2016) and Lang et al. (2014) using Massachusetts and Rhode Island data, respectively, and Hoen et al. (2015) examining wind farms across the country. In contrast, studies in European countries find that wind turbines have a significant negative impact on nearby properties, though the magnitude of the effect differs by region (Dröes and Koster, 2016, 2021; Gibbons, 2015; Jarvis, 2021; Sunak & Madlener, 2016). Vyn (2018) finds the Canadian experience to be heterogeneous and dependent on community acceptance. More recently, hedonic methods have focused on estimating externalities from offshore wind turbines. While this literature is still in its infancy, early studies indicate no negative impacts to property values or rental rates in the vicinity of offshore wind turbines (Jensen et al., 2018; Carr-Harris and Lang, 2019). Hedonic valuation has also been applied to residential rooftop solar. General consensus is that houses installed with rooftop photovoltaic panels sell for a premium, though there is regional variation in the size of the effect: 3.5% in California (Dastrup et al., 2012; Hoen et al., 2012), 5.4% in Hawaii (Wee, 2016), 17% in Arizona (Qiu et al. 2017), and 3.2% in Western Australia (Ma et al. 2016). However, this literature is only tangentially related as it is about quantifying internalities (e.g., valuation of personal financial benefits, warm glow), not externalities, and has nothing to do with land use.

property line setback. They survey residents of Rhode Island and find the largest determinant of approval is prior land use with positive WTP for arrays on non-greenfield sites ($10 to $21 per month per household) and negative WTP for arrays on greenfield sites (-$13 to -$49 per month per household). In addition, they find that WTP increases with size of the array and that households are willing to pay $6-8 per month to avoid full visibility. We contribute to this emergent literature by offering another data point in the understanding of externalities and providing a test of convergent validity for stated preference work.

Our work also relates to sociology and psychology research on renewable energy acceptance. Patrick Devine-Wright, a leading scholar in this area, argues in several articles against a simplistic NIMBY explanation, instead “local opposition is conceived as a form of place-protective action, which arises when new developments disrupt pre-existing emotional attachments and threaten place-related identity processes” (Devine-Wright, 2009). Our examination of heterogeneous treatment effects by prior land use and population density are consistent with the ideas of Devine-Wright and others (e.g., Brittan, 2001; Firestone et al., 2018; Wester-Herber, 2004) in that we are finding larger negatives associated with solar arrays developed in areas where this technology contradicts ‘local character’ and substantially alters the ‘positive distinctiveness’ that people associate with such places.. We contribute to this literature by adding a revealed preference, market-based metric of place meaning. The combination of frameworks seems particularly important when thinking about optimal siting of new infrastructure or prioritizing over multiple objectives in the energy landscape.

2 DATA

To implement the hedonic analysis, we build a composite dataset that integrates: 1) the data on the location and attributes of all solar developments in MA and RI, and 2) the data on attributes and locations of residential properties in MA and RI.

2.1 Solar data

The dataset on solar installations is obtained from the Energy Information Administration's (EIA's) report EIA-860M, or the Monthly Update to the Annual Electric Generator Report. The EIA-860M contains data on the total capacity of electric generation facilities in the United States that have a capacity of 1 MW and above, their point location (latitude and longitude), and the month and year that operation begins. Figure 1 represents a map of 284 solar installations constructed prior to August 2019, which is when we set the cutoff for being in our sample. The installations are well dispersed across all regions in both states, which increases confidence that estimates will not be affected by unobserved regional shocks. Figure 2 graphs new and cumulative solar capacity by year. The first installation began operation in June 2010. New capacity displays a continuous upward trend through 2014. There is a sharp fall in 2015, after which the trend rises again and peaks in 2017, before falling again in 2018. As of August 2019, the cumulative solar capacity of utility-scale arrays in MA and RI is 817 MW. Capacity factors for this region are about 16.5% (EIA 2019), which means these solar installations are collectively producing about 1180 GWh of electricity per year, which is enough to power 157,681 homes.

One limitation of our data is that we do not have shapefiles representing the exact footprint of the solar installations, thus we must approximate that using Geographic Information Systems (GIS) software. Solar installations require approximately 5 acres of land per MW of capacity (Denholm and Margolis, 2008; Ong et al., 2013). We assume that the point location is the centroid of the installation and then create a circle around it with an area equal to 5 times the capacity (in MW) of each array.[3]

We hypothesize that prior land use may affect property value impacts. Specifically, houses in proximity to farms and forests that are developed into solar may depreciate more than houses in proximity to a brownfield or capped landfill that

[3] We manually crosscheck the EIA data with Google Maps, and correct the latitude and longitude when they do not correspond to the centroid of the array. We recognize that this approximation of distance could lead some properties to be misclassified as treatment or control, inducing a small amount of measurement error in treatment status. As a result, our DID estimates may be slightly attenuated.

is developed into solar.[4] Since farms, forests, and other open space are amenities and boost home values (Irwin, 2002; Lang, 2018), conversion of these types of lands may lead to larger price decreases because it is the combination of a loss of amenities and the gain of disamenities. To infer prior land use, we overlay the estimated circular footprints on 2005 land use data obtained from Massachusetts Bureau of Geographic Information and 2011 land use data obtained from Rhode Island Geographic Information System for the respective states. We then assign each installation a prior land use: 'greenfield' if it was formerly either a farm or forest land, and 'non-greenfield' if it was either a commercial site or a landfill.[5] 63% of installations and 70% of capacity is classified as greenfield (see Figure A1 in the online appendix).

2.2 Property data

We use ZTRAX housing transaction data from Zillow (http://www.zillow.com/data), which include information on property location (latitude and longitude), sales price, date of transaction, and many property characteristics (lot size, square feet of living area, number of bedrooms, number of bathrooms, year built, number of fireplaces, central air-conditioning, and swimming pool). The data include 2,095,835 property transactions from January 2005 to June 2019 in the states of RI and MA. Condominiums and houses with missing observations for sales price, bedrooms, full bathrooms, and half bathrooms are dropped. We also drop groups of properties with the same latitudes and longitudes, but

---

[4] Solar developers prefer farm and forest lands because they have substantially lower construction costs compared to alternative sites like brownfields, covered landfills, parking lot canopies, and industrial areas.

[5] Several solar installations cover an area with multiple land uses. We obtain exactly one land use type per solar site in five additional steps. First, we classify the land use as 'landfill' if the installations have the term 'landfill' in their name, or if they are listed in the EPA's dataset of contaminated land. Second, we use a stratifying logic to group all land-use types under seven major categories: commercial, farm, forest, landfill, recreational, residential, and wetland. Third, we place *'transportation', 'urban public/institutional', 'industrial', 'powerline/utility'*, and *'junkyard'* under commercial; *'orchard', 'cropland', 'pasture', 'nursery',* and *'cranberry bog'* under farm; *'spectator recreation',* and *'participation recreation'* under recreation, *'multi-family residential', 'low density residential', 'medium density residential', 'very low density residential'*, and *'high density residential'* under residential; and *'forested wetland', 'water',* and *'non-forested wetland'* under wetland. Fourth, we rank all land use categories under each installation by area, such that the land use with the greatest area gets the highest rank. We drop all land use categories but the ones with the highest rank to obtain exactly one land use per installation in the following four major categories: commercial, farm, forest, and landfill.

different addresses because this indicates incorrect geocoordinates. Sales prices are adjusted to 2019 levels using the Northeast regional housing Consumer Price Index from Bureau of Labor Statistics. After dropping transactions with prices of $100 or less, since these are clearly not arms-length transactions, we drop transactions in the bottom and top 5% of the sales price distribution to get rid of outliers. Further, we drop observations that have more than four stories, six bedrooms, five full bathrooms, or three half bathrooms. Houses that underwent major reconstruction are dropped since they may have different attributes in previous transactions. We exclude homes that sell before they were built, as there is evidence these are lot sales without improved property. Properties that transact more than once on the same date are likely to be subdivisions and are therefore excluded. We also drop single-family residential properties with lot sizes larger than two acres, since large plots could be potential sites for solar development and price impacts of nearby solar could be completely different. Finally, we exclude all properties that transact only once in the chosen time frame because we focus on repeat sales analyses. We spatially merge the solar data with the property dataset by matching every property to the nearest eventual site of solar development to infer proximity.

Similar to prior land use, we hypothesize that existing development in areas surrounding solar arrays may impact property prices. Many rural areas pride themselves on their rural character and residents seek out that type of bucolic setting. Hence, construction of solar installations could be seen as an industrialization of the landscape and may cause larger negative impacts on property values. Whereas solar arrays in suburban and urban areas may be viewed as more congruent with existing surroundings. However, space is also more constrained in suburban and urban, which could lead to greater impacts there. We proxy for rural character with municipality-level population density, which comes from the 2010 Census. We define an indicator variable $Rural$, which equals one if the town has a population density of 850 people per square mile or fewer. We chose this cutoff because 850 is the average population density of MA, which forms the bulk of the observations in our dataset, and, at this cutoff, a little over a third of the properties and 66% of the solar installations are classified as rural, which we believe are reasonable proportions. It is important to note non-rural properties should not be thought of as urban, but more suburban. Very few utility-scale solar developments are built in urban areas as there is just not space.

3 METHODS

We use the difference-in-differences (DID) method in the hedonic framework to analyze the causal impact of solar installations on housing prices. We begin with a standard hedonic DID setup, in which we define treatment and control based on proximity. Properties located near large-scale solar installations are compared to similar properties that are further away from such installations, before and after construction.

We estimate repeat sales models that use within-property variation to identify the treatment effect by including property fixed effects:

$$P_{it} = \beta_1 Post_{it} + \beta_2 (Treated_i \times Post_{it}) + \boldsymbol{X_{it}\gamma} + \alpha_i + \epsilon_{it} \tag{1}$$

Where P_{it} is the log sale price of house i at time t. $Post_{it}$ is an indicator for post-treatment, which equals 1 if a house sells after the treatment date, and $Treated_i$ is a dummy variable equal to 1 if a house is located near an eventual solar site and 0 otherwise. $\boldsymbol{X_{it}}$ is a vector of temporal controls. α_i controls for time-invariant unobservables at the property level (e.g., school quality, proximity to other amenities and disamenities, traffic volume, walkability, property layout, curb appeal, etc.). Lastly, ϵ_{it} is the error term. In our basic specification, $\boldsymbol{X_{it}}$ includes month-year fixed effects, which capture macroeconomic trends that affect the entire region that could be correlated with solar development trends. In addition to this specification, we estimate two more models. The first adds municipality-specific time trends to account for different housing price trends between municipalities. The second includes county-year fixed effects which allows for county-specific, nonparametric differences in housing market trends. In all models, we cluster standard errors at the census tract level to allow for correlated errors within a larger area. β_1 is the change in prices for control properties from before to after treatment. β_2, the coefficient of interest, is the differential price change from before to after solar development for treated properties relative to control properties.

There are two aspects of this DID setup that are initially uncertain: the spatial extent of treatment, and the date on which treatment occurs. We define the treatment distance to be equal to 0.6 miles and provide evidence to support this choice in Section 3.1. Similarly, we specify the treatment date to be 6 months prior to when the solar array begins operating and provide supporting evidence in Section 3.2.

3.1 Spatial extent of treatment

Since the extent of treatment is unknown, we must identify *d*, the distance up to which the effects of constructing a solar installation persist, and this will define the boundary for our treatment group. Following similar strategies as Davis (2011), Muehlenbachs et al. (2015), and Boslett et al. (2019), we estimate a DID model similar to Equation (1), except with treatment defined in bins of successive tenth-mile increments and control always being 2-3 miles. Figure 3 plots the estimates for each tenth-mile distance bin ranging from zero to two miles. Results indicate large, negative impacts for houses within 0.1 mile, but with large standard errors. Point estimates are noisy, and some point estimates are close to zero. Bins 0.4-0.5 miles and 0.5-0.6 miles are negative and significant. Beyond 0.6 miles, all estimates are statistically insignificant. Given this evidence, in all future specifications, we define the treatment group to be within 0.6 miles and the control group to be 0.6 – 2 miles.[6]

We only include transactions occurring within two miles of any eventual solar installation to increase similarities in observable and unobservable characteristics for sample properties. For properties lying within 0.6 miles of two installations, we omit those that transact before the closer of the two installations is built, but after the further one is built. This removes only 0.04% of transactions and ensures a cleaner identification of the pre-treatment and post-treatment periods in our model.

3.2. Timing of treatment

The date on which treatment occurs in the minds of home buyers and sellers is ex ante unknown to us and is likely to pre-date the beginning of operation, which is

[6] Figure A2 in the online appendix plots the estimates from a similar regression, except with control defined as 1 – 2 miles. The results are qualitatively identical. Table A1 also examines robustness of results using different control groups based on different distances, and results are similar to the main findings.

the only milestone for which we have an exact date. To identify the treatment date, we conduct an event study that analyzes property price trends between the treated and control groups over time. Specifically, we define a time variable in terms of 6 month bins, starting from 6 years prior to operation date and up to 6 years post operation, and we choose 6 – 12 months prior as the reference category. We then estimate a DID model similar to Equation (1) in which we regress log sales prices on the treatment variable, the time bins, and their interaction, along with month-year and property fixed effects.

Figure 4 plots the coefficients and 95% confidence intervals of the event study model. There are two takeaways from these results. First, we find a large drop in prices in the 0-6 months prior bin relative to 6-18 months prior. The negative effect starting 6 months prior is sustained, though noisy, for the remainder of the post-operation period. As a result, we choose to define the treatment date as 6 months prior to operation date in all future specifications.[7] This timing is in line with our expectations because it takes time for the array to be constructed, and thus disamenities will be apparent to potential buyers prior to operation. Second, while noisy, there is evidence of a pre-treatment downward trend in prices, suggesting properties near eventual solar sites may have been declining prior to construction. This trend is punctuated by the large negative difference found in the 18-24 months prior to operation time period, but is then reversed in the 6-18 months prior periods. One or both could be anomalous, but the graph raises concerns about the viability of the necessary parallel trends assumption.[8] We discuss implications of this more in Sections 3.3-3.4.

3.3 Summary statistics and assumptions

[7] While we have presented versions of Figures 3 and 4 with the eventual spatial extent and treatment and treatment date included, these findings are robust to different choices of one or the other.
[8] Figure A3 in the online appendix presents a version of Figure 4 with time binned in increments of one year. There is less noise, but the qualitative findings that treatment begins six months prior to operation and relative prices are declining in treated areas pre-treatment hold.

Our final, composite dataset includes 107,291 repeat-sales transactions representing 45,795 unique properties around 282 solar installations.[9] We observe 11,292 transactions within 0.6 miles, of which 34% are post-treatment.

The summary statistics for key variables are given in Table 1. The first column represents the mean and standard deviation values of our full sample. The mean sales price is $314,710. The average property in our data has a lot size of 0.42 acres, has living area of just under 3000 square feet, approximately 3 bedrooms, and is about 58 years old. About 46% of the properties are matched to a greenfield development, and 35% are rural.

The critical assumption for the DID design to yield causal estimates is the parallel trends assumption, which requires that treatment and control properties would have the same trend in outcomes if treatment did not occur. We first assess the plausibility of this assumptions by comparing characteristics of treatment and control properties, with the logic that similar properties are likely to have similar price trajectories. The second and third columns in Table 1 compare pre-treatment housing attribute means between the 0 – 0.6 miles (treated) and 0.6 – 2 miles (control) observations. In the fourth column, we report the differences in means and their standard errors, which are estimated by regressing each housing attribute on treatment status in the pre-treatment period, along with month-year fixed effects. Several variables have a statistically significant difference, which raises concerns over the comparability of the control group. In the final column, we estimate differences in means conditioning on census block fixed effects. In this case, none of the housing attributes have a statistically significant difference in pre-treatment means, suggesting that the addition of spatial controls mitigates covariate imbalance. Our regression model uses property fixed effects, which effectively removes any concern about covariate overlap, except if price trends are correlated with housing characteristics.

Second, we examine pre-treatment trends in sales prices as seen in Figure 4. As discussed above, most coefficients hover near zero and are statistically insignificant in the pre-treatment period. However, the coefficient for 18-24 months prior to operation

[9] Our original dataset had 284 solar installations, but two are dropped because there are no repeat sales properties within 2 miles of them.

is negative, significant and large in magnitude. Additionally, there is some evidence of an overall negative trend in the pre-treatment coefficients, though the coefficients 6-12 and 12-18 months prior run counter to that trend. Thus, the evidence is not convincing either in support or refutation of the parallel trends assumption. Our identification strategy detailed in Equation (1) will mitigate bias from unobserved, time-invariant factors that are correlated with housing prices and solar siting. However, if the precise location of a solar array is endogenous and correlated not just with time-invariant unobserved attributes, but also correlated with price trends, then a comparison of treatment to control areas may be biased. Therefore, in the following section, we discuss an alternate DID estimator we employ that does not rely on a non-proximate control group and thus removes bias stemming from site selection being correlated with price trends.

3.4 Alternative DID estimator

We consider an alternative DID design that does not rely on a never-treated control group. We estimate Equation (1) on a subsample that includes only treated observations, and drops all properties that lie greater than 0.6 miles away from the nearest eventual solar installation. Identification in this model relies entirely on temporal variation in the construction of solar installations, instead of a combination of this variation and variation in trends between near and far houses. One inconsequential change is that the variable $Post_{it}$ is collinear with $Treated_i \times Post_{it}$ and drops out from the model. Our coefficient of interest is still β_2 and has the same interpretation. We only estimate these models including either municipality-specific time trends or county-year fixed effects since there is no spatially proximate never-treated control group to capture local time trends.

Figure 5 presents a pre-treatment price trends analysis for this alternative estimator. Examining Figure 2, we essentially divide the sample in half based on operation year, the thought being that properties near solar arrays built later can serve as a control for properties near solar arrays built earlier. We exclude properties near solar arrays built in 2010 and 2011 because this will allow more pre-treatment years to be examined and we lose relatively few observations by doing so. We define treatment as an indicator variable equal to 1 if a house is proximate to a solar installation that will be built in years 2012 – 2016 and equal to zero if proximate to a solar installation that will be built in 2017 or later. Log sales prices are regressed on the interaction between treatment and year dummy variables, along with month-year, property, and county-year fixed effects. Estimated coefficients giving differences between treatment and control properties over time and 95% confidence intervals are graphed. The price trends look similar, with no evidence of any downward trend as in Figure 4.

DID methods excluding the never-treated group have been applied before in many settings for myriad reasons. As mentioned in the introduction, Dröes and Koster (2021) use this approach in their hedonic study of solar arrays and wind turbines in the Netherlands out of concern for endogeneity of siting decisions. Beatty et al. (2021) only include treated gas stations in their preferred model of price impacts of hurricanes due to concerns about SUTVA violations. Lang and Cavanagh (2018) only include treated properties in their hedonic study of brownfield remediation because the density of brownfields made never-treated controls not proximate to treated observations and housing characteristics were dissimilar.

In Section 4, for every specification, we present results using both the DID model that includes the never-treated properties and the DID model that excludes the never-treated properties.[10] We remain equivocal about which is preferred and instead focus on the range of estimates.

3.5 Heterogeneity in treatment effect

We extend the analysis to investigate heterogeneity in treatment effect in multiple ways. First, we investigate heterogeneity in treatment effect by two place-based characteristics: prior land use and rural character. This is done by a triple

---

[10] An additional concern with staggered DID models is that estimated coefficients can be biased if treatment effects are heterogeneous over time and some observations have negative weights (de Chaisemartin and D’Haultfœuille, 2020). We analyze our data for the presence of negative weights and find relatively few. Applying the twowayfeweights command in Stata (de Chaisemartin et al., 2019), 12.6% of treated observations have an associated negative weight and the sum of negative weights is -0.0048. This compares favorably to the case study data used by de Chaisemartin and d'Haultfoeuille (2020) in which 40.1% of treated observations have an associated negative weight and the sum of negative weight is -0.533. Further, the time corrected wald estimator proposed by de Chaisemartin and D’Haultfœuille (2020) produces estimates qualitatively identical to the standard DID. Thus, we are not concerned about this particular source of bias.

difference analysis in which we interact the treatment effect term in Equation (1) with variables for our characteristic of interest. The specifications are as follow:

$$P_{it} = \beta_1(Post_{it} \times Non-greenfield_i) + \beta_2(Treated_i \times Post_{it} \times Non-greenfield) + \beta_3(Post_{it} \times Greenfield_i) + \beta_4(Treated_i \times Post_{it} \times Greenfield_i) + \boldsymbol{X_{it}\gamma} + \alpha_i + \epsilon_{it} \quad (2)$$

$$P_{it} = \beta_1(Post_{it} \times Non-rural_i) + \beta_2(Treated_i \times Post_{it} \times Non-rural) + \beta_3(Post_{it} \times Rural_i) + \beta_4(Treated_i \times Post_{it} \times Rural_i) + \boldsymbol{X_{it}\gamma} + \alpha_i + \epsilon_{it} \quad (3)$$

where $Greenfield_i$ is an indicator variable equal to 1 if a property is located within the vicinity of a solar installation that was built on land that was formerly farmland or forested and $Non-greenfield_i = 1 - Greenfield_i$. $Rural_i$ is an indicator variable equal to 1 if property i lies in a town with a population density of 850 people per square mile or fewer, and $Non-rural_i = 1 - Rural_i$.

Our coefficients of interest in Equations (2) and (3) are β_2 and β_4. In Equation (2), we hypothesize that $\beta_4 < \beta_2 < 0$ because developments on farm and forest lands will lead to larger negative impacts on housing prices due to the more dramatic change in landscape compared to a commercial site or landfill and the loss of open space amenities. In Equation (3), we again hypothesize that $\beta_4 < \beta_2 < 0$ because solar arrays are less congruent with rural settings and the contrast will lead to greater price declines, but there's more uncertainty here because of land scarcity in non-rural areas.

Second, we estimate a model that allows for heterogeneity in the impact based on distance. We identified treatment extending to 0.6 miles in Figure 3, but Figure 3 also suggests that treatment effects could be larger within 0.1 mile. To explore this possibility more formally, we develop a model that defines multiple distance bands. The first (outermost) band represents control properties located 1 – 2 miles away from the nearest solar installation. The second band is properties 0.6 – 1 mile away, which we differentiate from 1 – 2 miles to further test if the spatial extent of treatment does end at 0.6 miles. The third band includes treated properties located 0.1 – 0.6 miles from the nearest solar installation. Finally, the fourth (innermost) band consists of treated properties within a distance of 0.1 mile from the closest installation. Our specification is:

$$P_{it} = \beta_1 Post_{it} + \sum_{k=2}^{4} \beta_2^k \left(dist_i^k \times Post_{it}\right) + \boldsymbol{X_{it}\gamma} + \alpha_i + \epsilon_{it} \tag{4}$$

where $dist_i^k$ is a dummy variable equal to 1 if a property i lies within the k^{th} distance band. P_{it}, $Post_{it}$, X_{it}, and α_i are as defined in Equation (1). When estimating this model excluding never-treated properties, we only get estimates on the two inner rings, 0-0.1 miles and 0.1-0.6 miles.

4 RESULTS

4.1 Average treatment effects

We present our results estimating Equation (1) in Table 2. Columns 1 – 3 include the never-treated control group (distances of 0.6 – 2 miles), while Columns 4 and 5 exclude the never-treated properties. All columns include month-year fixed effects and property fixed effects, Columns 2 and 4 additionally include municipality-year time trends, and Columns 3 and 5 replace those with county-year fixed effects. Including never-treated properties yields treatment effect coefficient estimates that range from -0.015 to -0.024. Excluding the never-treated properties yields coefficients that are about twice as large, ranging from -0.028 to -0.036. The smaller magnitudes observed in Columns 1 – 3 likely stem from the pre-treatment, downward trend in treated properties relative to never-treated properties seen in Figure 4. Overall, treatment effects are negative and statistically significant across all models, confirming our hypothesis that nearby solar installations are, on average, a disamenity. Estimates suggest that houses lying within 0.6 miles of solar installations sell between 1.5% and 3.6% less post construction, all else equal.

We convert the percentage reduction to dollars by multiplying the coefficient and the average, pre-treatment property price for treated properties ($314,710), which gives us a range of $4,721 - $11,330. Assuming capitalization can be converted to a welfare measure in this context (see Kuminoff and Pope, 2014), we can then translate this price discount into an annual willingness to pay for avoiding proximity to solar. Assuming a 5% interest rate, average annual willingness to pay is $236 - $567 per household.

4.2 Heterogeneous treatment effects

In Table 3, we examine the heterogeneity in treatment effect by three characteristics: prior land use, rural character of towns, and proximity to solar installations. Each panel presents two specifications, mirroring the sample and control variables in Columns 3 and 5 of Table 2.

In Panel A, we provide estimates from the model described by Equation (2) where we explore heterogeneity by prior land use. The results conform to our expectations; estimated treatment effects for greenfield and non-greenfield sites are both negative, but the treatment effects for greenfield sites is larger in magnitude. The coefficient on $Treated \times Post \times Greenfield$ ranges from -0.020 to -0.044 and is significant at the 5% level or higher in both specifications. In contrast, the coefficient on $Treated \times Post \times Non - greenfield$ ranges from -0.011 to -0.013 and is not statistically significant in either column.

There are two questions that arise from these results. First, do non-greenfield sites have zero externalities? Statistically, yes, we fail to reject a null hypothesis of no effect. However, the coefficients are consistently negative (also holds in Table 4 discussed below), so there may be some signal there, just not enough to overcome the noise. Additionally, across non-greenfield sites, there could be additional heterogeneity that we are unable to measure. For example, different arrays could have varying degrees of visibility. To improve our understanding, we can draw on Lang et al. (2021), who recently conducted a choice experiment survey on preferences for solar siting attributes in Rhode Island. They estimate separate models for greenfield and non-greenfield solar sites and find that respondents have positive WTP to avoid fully visible arrays for both types ($10.34/month for greenfields and $4.42/month for non-greenfields).[11] In addition, respondents prefer further setback from property boundaries for non-greenfield sites, but are indifferent about setback on greenfield sites. Thus, these choice experiment results indicate that negative externalities can be

[11] This is not to be confused with Lang et al. (2021)'s finding of a positive total WTP value for siting arrays on non-greenfield land types. This result is not related to site proximity and instead reflects overall preferences for siting solar arrays on non-greenfield land types.

present at non-greenfield sites. We argue that our hedonic estimates reflect those negative externalities, though we cannot be confident in the exact magnitude of those effects.

The second question raised by the results of Panel A Table 3 is whether the difference of the greenfield treatment effect relative to the non-greenfield treatment effect is driven entirely by loss of open space. This is a critical question because if the alternative to solar arrays is residential housing and that leads to the same disamenities, then there is no reason to be concerned about solar developments. In truth, we cannot definitively know, but we argue there are attributes of a solar array that lead to additional negative externalities beyond residential development of open space. Some portion of the wedge could be due to nearby residents feeling that solar arrays are incongruent with that type of landscape and it takes away from the aesthetic of that place in ways that common houses do not. Additionally, we can again point to the Lang et al. (2021) results that show greater negative viewshed externalities on greenfield sites relative to non-greenfield sites, and these are on top of already substantial WTP to avoid development on greenfield sties to begin with. Relatedly, an unintended byproduct of this analysis is providing an upper bound on the value of privately held open space. Irwin (2002) and Geoghegan et al. (2006) both examine the property value impacts of developable open space (as well as permanently conserved open space), but they use a cross sectional approach, and their estimates vary substantially across models with some indicating developable open space is valued more than residential development and some the opposite. Our research offers better identification and bounds the impacts of loss of nearby developable open space as a small negative.

In Panel B, we examine heterogeneity by rural character of towns and report the coefficients from the specification defined in Equation (3). Similarly, these results conform to our expectations; estimated treatment effects for rural and non-rural sites are both negative and the treatment effect for rural sites is on average larger in magnitude. The coefficient on $Treated \times Post \times Rural$ ranges from -0.025 to -0.058 and is significant at the 5% level or higher in both columns. In contrast, the coefficient on $Treated \times Post \times Non-rural$ ranges from -0.005 to -0.006 and is not statistically significant. The results suggest that nearby utility-scale solar causes housing prices to decline more in rural areas than suburban or urban areas.[12]

There is, as expected, a strong positive correlation (0.41) between greenfield and rural, which raises the question of whether large negative results observed in rural areas are just a function of the higher proportion of greenfields found there or vice versa. Table A3 in the online appendix estimates a quadruple interaction model to try to parse the effects of greenfields and arrays in rural areas. It is clear that the smallest impacts accrue to properties near non-rural, non-greenfield sites. However, other orderings are inconsistent across columns, with each of three other categories yielding the largest negative impact in at least one specification. These results suggest we cannot attribute the results to greenfield sites or rural sites alone. In addition, these results bolster our claim that greenfield treatment effects are not entirely due to loss of open space. In sum, the results of Panels A and B indicate that valuation depends on context; surrounding land uses and place meaning contribute to the magnitude of price declines.

Lastly, in Panel C, we estimate the model described by Equation (4) that allows for heterogeneity in the impact on prices based on distance. The coefficient for the 0.6 – 1 mile band is statistically insignificant in Column 1, which is consistent with our assumption that treatment effects do not persist beyond 0.6 miles. The coefficients on the 0.1 – 0.6 mile band are significant and similar magnitude to the main results. The coefficients on the 0 – 0.1 mile band range between -0.038 to -0.042, which is between 1.5 to 2.4 times larger in magnitude than the 0.1 – 0.6 mile band, though insignificant. These results are suggestive that property values for homes lying within 0.1 mile from a solar installation may fall substantially, but our estimates are imprecise reflecting few observations within that distance band.[13]

In the online appendix, we also present results that test for heterogeneity by size of installation and time since construction (see Tables A4 and A5). We find that there are no statistically significant differences between categories, and results suggest

---

[12] We examine different population density cutoffs for the definition of Rural in Table A2 in the online appendix. Results are consistent across different cutoffs.
[13] There are 218 observations lying within 0.1 miles from a solar installation, of which 72 sell post construction.

that larger installations do not cause greater price declines and that treatment effects do not dissipate with time.

4.3 Robustness checks

In Table 4 we present results from a series of robustness checks to ensure that our results are consistent to alternative data constructions or samples. We present results both for the average treatment effect models and models focused on greenfield heterogeneity. We include the latter because it is a critical piece of the story. Further, we present results for both the models that include the never-treated control properties (Panels A and B) and exclude the never-treated properties (Panels C and D) do the same, except using the ever-treated sample.

Columns 1 and 2 explore the assumption of the amount of land required per MW of installed capacity. Instead of 5 acres in our main models, Column 1 assumes 4 acres, and Column 2 assumes 6 acres. By contracting and expanding the assumed size of installations, the set of properties that are designated as treatment and control are altered. The estimated coefficients in these columns are qualitatively identical to the main and heterogeneity results, indicating that assumptions about the radius of arrays is not impacting results.

In Column 3, we control for the presence of wind turbines by including an indicator variable equal to 1 if a house lies within one mile of a built wind turbine. One may be concerned that solar and wind are co-located and disamenities from one may be captured in the estimated valuation of the other if not controlled for. The treatment coefficient is nearly identical to the main results. In MA and RI, there is little correlation in the siting of wind and solar energy, and solar is far more abundant (see Figure A4 in the online appendix).

Our main sample includes transactions in years 2005-2019. One may be concerned that this is too long of a time horizon and changes to the hedonic function can occur over that time. To address this concern, Column 4 only includes transactions occurring 2009-2019, and Column 5 only includes transactions that are within four years before or after the treatment date of the solar installation they are matched to. Both of these sample restrictions, particularly Column 5, greatly reduce our sample size in the repeat sales model because fewer properties transact multiple times in a short window. The average treatment effect estimates are larger, and the greenfield treatment effects are over twice as large.

In the online appendix, we check the robustness of our main results in three more ways. First, in Table A6 we test for anticipation effects two years prior to solar farm operation date and find no evidence of anticipation. Second, for the model that includes the never-treated properties, we vary the spatial extent of the control group (Table A1) and find that the treatment effect is robust to different control group boundaries. Finally, in Table A7 we examine whether regional price trends may be correlated with solar installation construction by including distance to city center by year trends in all our specifications. Our coefficients remain robust, suggesting that this is not a threat to identification.

5 Conclusion

This paper estimates the valuation of externalities associated with nearby utility-scale solar installations using revealed preferences from the property market. Using the DID empirical technique, we define treatment by distance to the nearest solar installation and compare treated properties to those lying between 0.6 and 2 miles from the installation (never-treated group), or to properties that receive a solar installation in their vicinity in the future. We observe 11,292 housing transactions occurring within 0.6 miles (treated group), and 95,999 transactions between 0.6 and two miles (never-treated control group) of 282 solar installations in MA and RI.

Our findings can be summarized as follows: there is a consistent negative average effect of proximity to utility-scale solar array, the estimates derived using the ever-treated sample are consistently larger than the ones that use the never-treated control group, and arrays sited on greenfields an in rural areas cause larger negative impacts and drive the overall negative and significant average effects. Average treatment effects suggests that property values decline between 1.5% ($4,721) and 3.6% ($11,330) after the construction of a nearby solar installation, all else equal. This translates to an annual willingness to pay between $236 and $567 per household to avoid disamenities associated with proximity to the installations.

While a full benefit-cost analysis (BCA) of utility-scale solar arrays is beyond the scope of this paper, because we do not know anything about consumer and producer surplus[14], we can at least benchmark the negative, local externalities against the global benefits of greenhouse gas (GHG) mitigation. We therefore conduct the following back-of-the-envelope calculations. While solar arrays typically have a lifetime of 25-30 years, there is uncertainty about what would happen after that time. Thus, we ignore those dynamic issues and only calculate costs and benefits for a single year. On the cost side, we first consider the point estimate from our preferred specifications, which translate to a loss between $236 and $567 per year per household for treated homes close to solar installations. Our complete dataset (prior to any sample cuts) consists of 72,538 unique properties located within 0.6 miles of all solar installations in the dataset. Put together, we estimate an annual welfare loss between $17.12 and 41.13 million due to proximate solar installations in MA and RI.

To quantify the GHG benefits from solar installations, we first calculate net generation from solar installations. Assuming a capacity factor of 16.5%, the 817 MW of installed solar capacity in MA and RI generates is 1,180,892 MWh (megawatt hours) of electricity per year.[15] Current non-renewable generation in MA and RI comes almost entirely from natural gas. According to the EIA, 0.42 mt (metric tons) of CO_2 are emitted from each MWh of electricity that is generated from natural gas, implying that a total of 495,975 mt of CO_2 are abated annually from solar energy generation. In addition, natural gas can leak in the distribution system, which releases methane, a much more potent greenhouse gas. Based on Hausman and Muehlenbachs (2018) and EIA, each MWh generated from natural gas is associated with 104.72 cf (cubic feet) of methane leaked. Thus, one year of solar generation mitigates an estimated 123,663 mcf of leaked methane, which has equivalent warming potential of 48,538 mt of CO_2. The EPA estimates the current social cost of CO_2 is $51.80 per metric ton, which places the value of annual greenhouse gas mitigation to be $28.21 million (US EPA).

[14] To be sure, significant amounts of money are part of the market transactions. A developer quoted us that they offer landowners $15-20,000 per MW per year of installed capacity. It is unknown how much is profit and whether some portion of that could be used to compensate proximate households.

[15] $Net\ generation\ (MWh) = \%\ Capacity factor \times 365\ days \times 24\ hours \times Installed\ capacity\ (MW)$

Combining the estimates, the benefit-cost ratio is between 1.65 and 0.69. In one scenario (using the DID estimates including the never-treated properties), the global benefits outweigh the local costs. However, using the DID estimates that exclude the never-treated properties, we come to the opposite conclusion. Regardless, in both cases it is clear that the local costs are substantial, bolstering local concerns about solar siting and clarifying the magnitude of costs borne by neighboring property owners. However, the benefit-cost ratio may be substantially better in other states that are less densely populated and more reliant on coal.

This research offers policy relevant findings. Communities in southern New England and elsewhere in the United States are currently grappling with contentious solar siting issues and will be for some time. These results quantify some of the opposition to certain siting decisions and allow those voices to enter into a state or local BCA. Further, our results suggest ways to reduce negative externalities that could be activated by state and local governments. In the case of siting on brownfields and covered landfills, developers may require additional subsidies to target those areas. Though non-financial costs, such as faster zoning approval may compensate them as well.

There are several directions of important future research. Similar hedonic studies should be completed elsewhere in the United States to assess similarity of valuation estimates and test our assertion that benefit-cost ratios will be more favorable elsewhere. Though, as discussed above, Abashidze (2019) finds even larger impacts than ours in North Carolina. In addition, examining valuation of smaller solar arrays (100 kW – 500 kW) could yield new insights. In southern New England, farms can install arrays of this size on marginal land and generate income that can help sustain the farm in the face of rising land costs (EcoRI, 2020). The main policy recommendation stemming from our analysis is to move solar development away from greenfields and away from rural areas if the objective is to minimize externalities. However, implementing such a strategy could have unintended consequences for environmental justice because it is marginalized communities that live in urban areas and nearby non-greenfield lands. Future research can better inform how the construction of solar installations differentially impacts minorities and low income communities. Lastly, community solar is a popular idea that is understudied in the context of siting. A contingent valuation survey could assess willingness to accept proximate community solar if it was structured such that nearby residents benefited financially.

Figures and Tables

Figure 1: Map of utility-scale solar installations across Massachusetts and Rhode Island

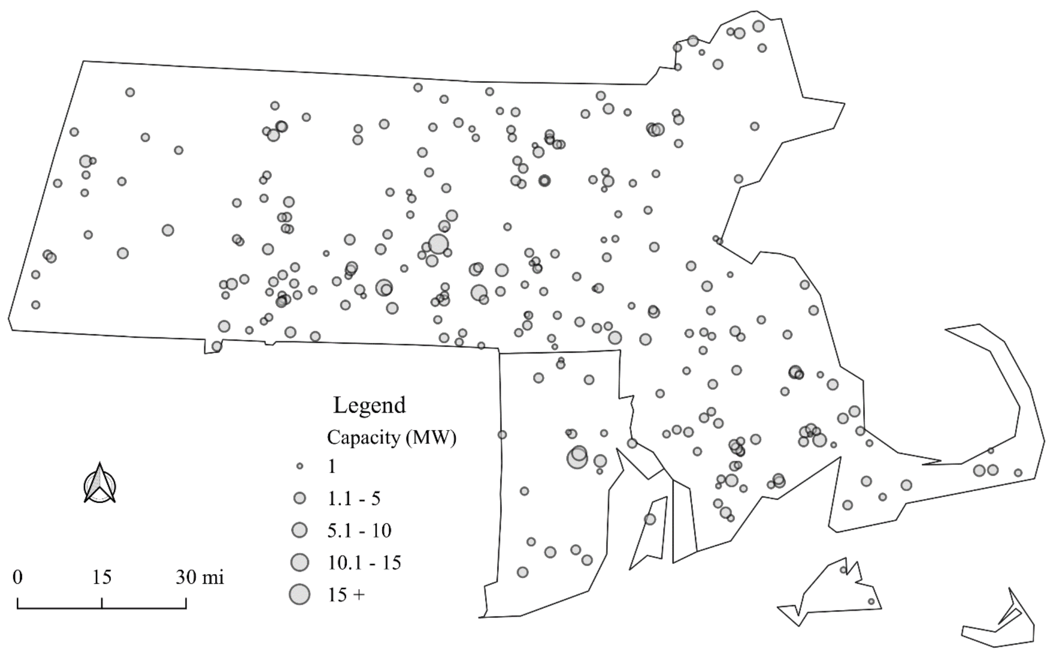


Figure 2: New and cumulative utility-scale solar capacity by year

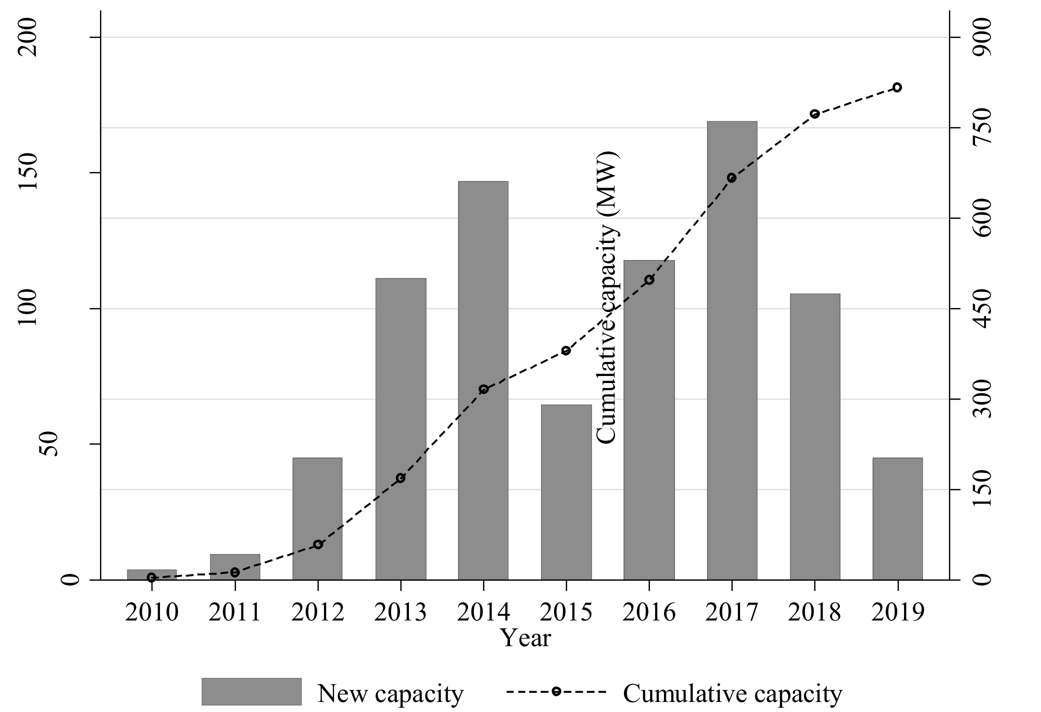
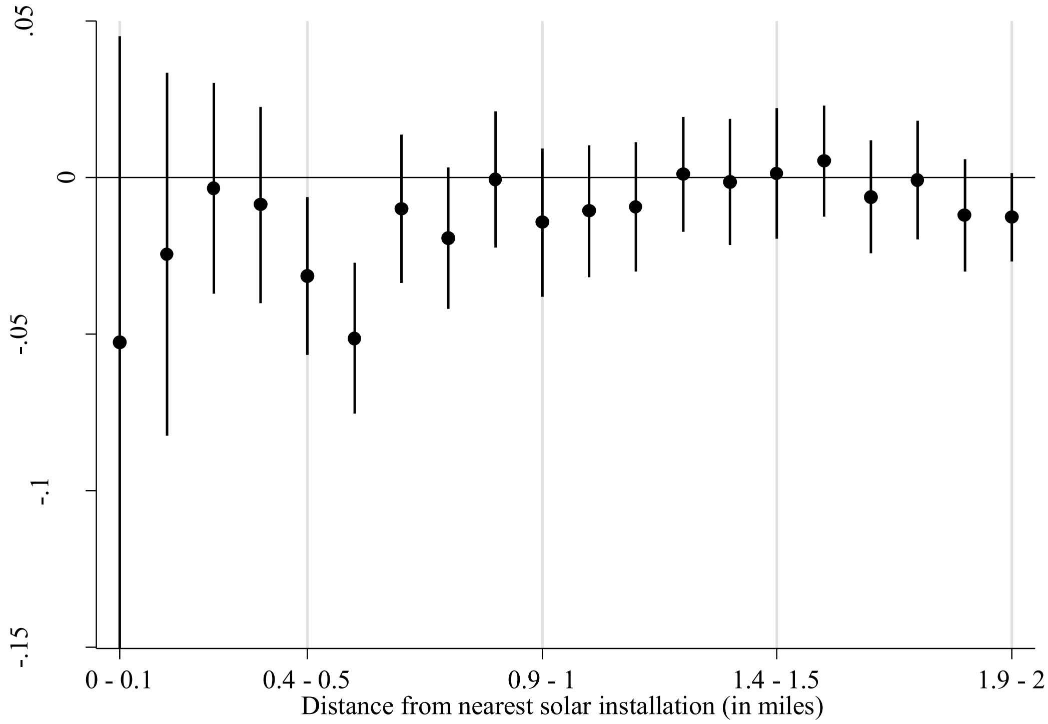
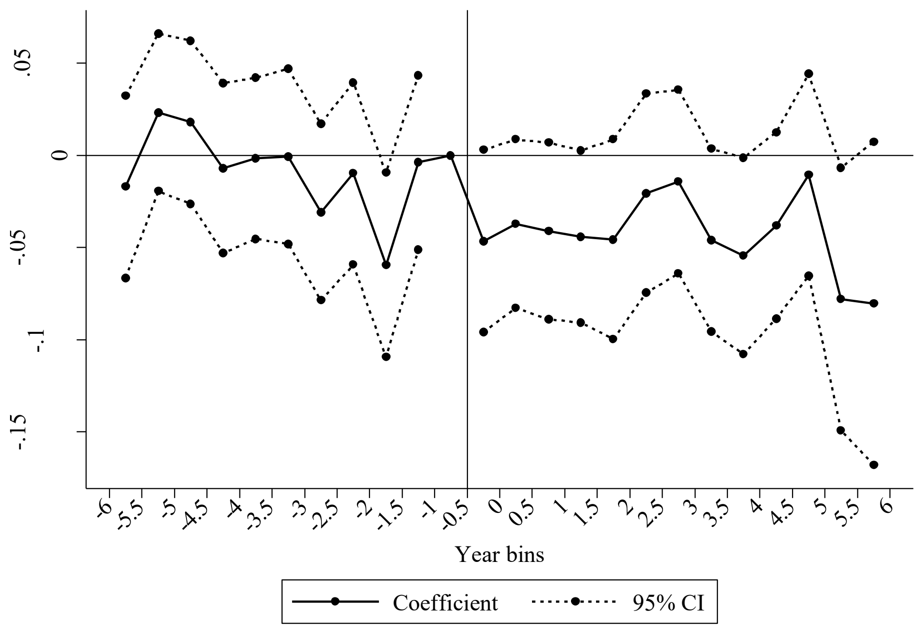


Figure 3: Spatial extent of treatment



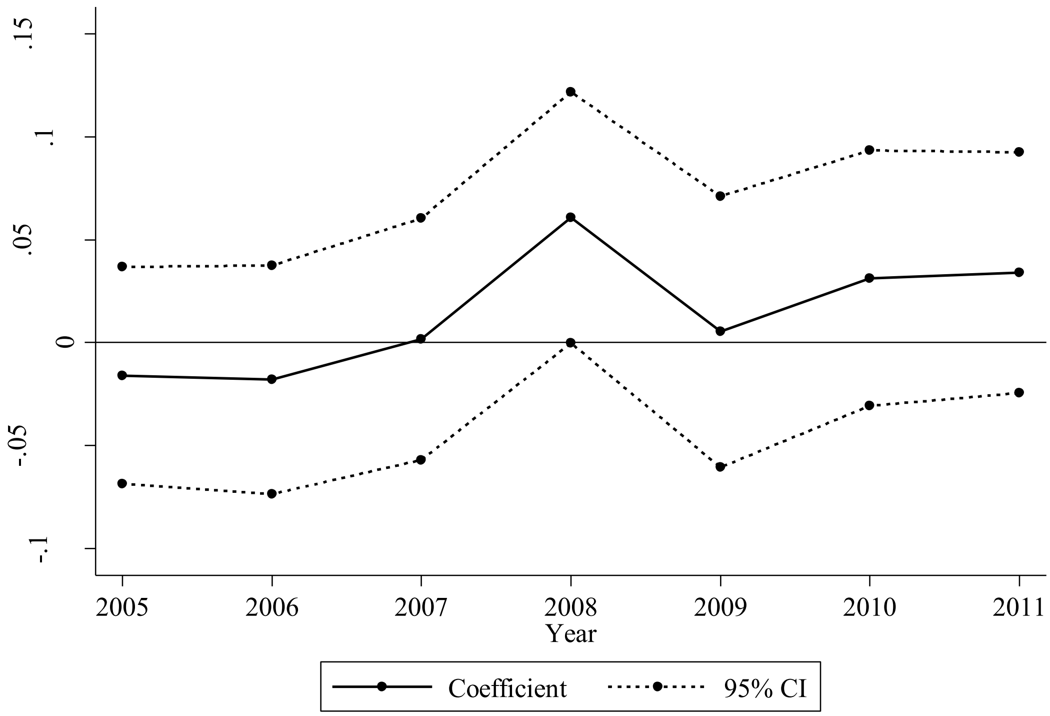
Notes: The treatment variable is defined as a bin variable, with treated properties lying within 1/10 mile distance bands up to 2 miles. Control properties are those lying 2 – 3 miles away from the nearest solar installation. Post = 1 if a house sells after the treatment date. The treatment date is defined as 6 months prior to solar installation operation date. The coefficients are obtained by estimating a DID model similar to Equation 1 that regresses log sales price on 1/10 mile distance bands up to 2 miles, along with month-year and property fixed effects. Resulting coefficients and 95% confidence intervals are graphed.

Figure 4: Event study of prices before and after solar installation operation date



Notes: The treatment variable is defined as a dummy variable equal to 1 if a house is within 0.6 miles of an eventual solar installation site. The time period variable is defined as a bin variable, starting from 6 years prior to solar installation operation date and up to 6 years post operation. Properties are sorted into the respective 6 month bin in which they transact, and the reference time period is 0.5 to 1 year prior to operation date. The coefficients are obtained by estimating a DID model similar to Equation 1 that regresses log sales price on the interaction between the treatment and the time period variables, along with month-year and property fixed effects. Resulting coefficients and 95% confidence intervals are graphed.

Figure 5: Pre-treatment price trends for DID model excluding never-treated properties



Notes: Sample size is 10,452 and includes properties within 0.6 miles of an eventual solar site built 2012 or later. Treatment is defined as an indicator variable equal to 1 if a house is proximate to a solar installation that will be built between 2012 and 2016 and equal to 0 if proximate to a solar installation built in 2017 or later. Log sales prices are regressed on the interaction between treatment and year dummy variables, along with month-year, property, and county-year fixed effects. Estimated coefficients giving differences between treatment and control properties over time and 95% confidence intervals are graphed.

Table 1: Housing attribute means by treatment status

Variable	Full sample means (std. dev.)	Pre-treatment means (std. dev.) 0 - 0.6 miles	Pre-treatment means (std. dev.) 0.6 - 2 miles	Difference in means (std. error)	Difference in means with spatial controls (std. error)
Price (000's)	314.71	305.36	316.37	-11.630	-1.907
	(159.73)	(158.08)	(160.45)	(8.688)	(6.075)
Lot size (acres)	0.42	0.47	0.41	0.057***	0.008
	(0.37)	(0.40)	(0.37)	(0.020)	(0.019)
House area (000's sq. feet)	2.92	2.86	2.92	-0.064	0.045
	(1.22)	(1.26)	(1.22)	(0.072)	(0.061)
Bedrooms	3.09	3.04	3.09	-0.046**	-0.015
	(0.71)	(0.70)	(0.72)	(0.019)	(0.019)
Full bathrooms	1.53	1.52	1.53	-0.010	-0.016
	(0.62)	(0.62)	(0.62)	(0.024)	(0.032)
Half bathrooms	0.48	0.45	0.48	-0.032**	-0.019
	(0.52)	(0.52)	(0.52)	(0.016)	(0.025)
Age of home (years)	58.06	50.19	56.08	-5.888***	1.487
	(35.71)	(33.26)	(35.75)	(1.526)	(1.785)
Pool (1 = yes)	0.05	0.05	0.05	-0.002	0.004
	(0.21)	(0.22)	(0.22)	(0.008)	(0.011)
Air conditioning (1 = yes)	0.35	0.38	0.35	0.039**	0.002
	(0.48)	(0.49)	(0.48)	(0.020)	(0.022)
Fireplace number	0.39	0.35	0.40	-0.047**	0.000
	(0.58)	(0.56)	(0.58)	(0.023)	(0.026)
Greenfield (1 = yes)	0.46	0.49	0.47	0.024	-0.002
	(0.50)	(0.50)	(0.50)	(0.036)	(0.001)
Rural (1 = yes)	0.35	0.39	0.37	0.049	-0.003
	(0.48)	(0.49)	(0.48)	(0.034)	(0.003)
Observations	107,291	7,448	64,322		

Notes: Sales prices are adjusted to 2019 levels using the CPI. In Column 4, the differences in means and standard errors are estimated by regressing each housing attribute on treatment status and month-year fixed effects, using only pre-treatment transactions. Column 5 adds census block fixed effects to the regression model used in Column 4 and reports the estimated differences in means and standard errors. Standard errors are clustered at the tract level. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Table 2: Estimates of the impact of solar installations on property prices

Independent variables	Dependent variable: Sale price (ln)				
	Include never-treated			Exclude never-treated	
	(1)	(2)	(3)	(4)	(5)
Post	0.017***	-0.003	-0.005		
	(0.005)	(0.005)	(0.005)		
Treated × Post	-0.024***	-0.018***	-0.015**	-0.036***	-0.028**
	(0.008)	(0.007)	(0.007)	(0.013)	(0.013)
Controls					
Month-year fixed effects	Y	Y	Y	Y	Y
Property fixed effects	Y	Y	Y	Y	Y
Municipality time trends		Y		Y	
County-year fixed effects			Y		Y
Observations	107,291	107,291	107,291	11,292	11,292
R^2	0.871	0.876	0.878	0.889	0.891

Notes: Treated = 1 if a house is within 0.6 miles of an eventual solar installation site and Post = 1 if a house sells after the treatment date. The treatment date is defined as 6 months prior to solar installation operation date. Standard errors are clustered at the tract level and shown in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Table 3: Heterogeneity of treatment effects

Independent variables	Dependent variable: Sale price (ln)	
	(1)	(2)
Panel A: Heterogeneity by prior land use		
Treated × Post × Non-greenfield	-0.011	-0.013
	(0.010)	(0.015)
Treated × Post × Greenfield	-0.020**	-0.044***
	(0.009)	(0.014)
Panel B: Heterogeneity by population density		
Treated × Post × Non-rural	-0.006	-0.005
	(0.009)	(0.015)
Treated × Post × Rural	-0.025**	-0.058***
	(0.010)	(0.014)
Panel C: Heterogeneity by proximity		
(0.6 – 1 mile) × Post	-0.005	
	(0.006)	
(0.1 – 0.6 miles) × Post	-0.016**	-0.028**
	(0.007)	(0.013)
(0 – 0.1 miles) × Post	-0.038	-0.042
	(0.052)	(0.053)
Observations	107,291	11,292

Notes: Both specifications include property, month-year, and county-year fixed effects. Treated = 1 if a house is within 0.6 miles of a solar construction and Post = 1 if a house sells after the treatment date. The treatment date is defined as 6 months prior to solar installation operation date. Greenfield = 1 if the prior land use is farm or forest land and Non-greenfield = (1-Greenfield). Rural = 1 if the population density per square mile is ≤ 850 and Non-rural = (1-Rural). In column 1, the Panel A model also includes Post × Greenfield and the Panel B model includes Post × Rural, though neither coefficient is displayed. In Panel C, (0.6 – 1 mile), (0.1 – 0.6 miles), and (0 – 0.1 mile) are dummy variables = 1 if properties lie within the respective distances from the nearest eventual solar installation, and distance bin for 1 – 2 miles is omitted in Column 1. Standard errors are clustered at the tract level and shown in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Table 4: Robustness checks

Independent variables	1 MW = 4 acres (1)	1 MW = 6 acres (2)	Control for wind turbines (3)	Drop pre-2009 transactions (4)	Keep properties transacting +/- 4 years from treatment date (5)
	Dependent variable: Sale price (ln)				
	Include never-treated				
Panel A: Standard model					
Treated × Post	-0.014*	-0.014**	-0.015**	-0.025***	-0.021*
	(0.007)	(0.007)	(0.007)	(0.009)	(0.012)
Panel B: Heterogeneity model					
Treated × Post × Non-greenfield	-0.011	-0.009	-0.011	-0.014	-0.003
	(0.010)	(0.010)	(0.010)	(0.013)	(0.016)
Treated × Post × Greenfield	-0.017*	-0.019**	-0.020**	-0.037***	-0.043***
	(0.009)	(0.009)	(0.009)	(0.012)	(0.016)
	Exclude never-treated				
Panel C: Standard model					
Treated × Post	-0.024*	-0.030**	-0.027**	-0.038**	-0.066**
	(0.013)	(0.013)	(0.013)	(0.016)	(0.028)
Panel D: Heterogeneity model					
Treated × Post × Non-greenfield	-0.009	-0.015	-0.011	-0.010	-0.027
	(0.016)	(0.015)	(0.015)	(0.019)	(0.031)
Treated × Post × Greenfield	-0.038***	-0.045***	-0.044***	-0.067***	-0.111***
	(0.014)	(0.014)	(0.014)	(0.017)	(0.032)

Notes: Treated = 1 if a house is within 0.6 miles of a solar construction, and Post = 1 if a house sells after the treatment date. The treatment date is defined as 6 months prior to solar installation operation date. In Columns 1 and 2 we assume that the land area required for 1 MW of solar capacity is 4 acres and 6 acres, respectively. We control for the presence of wind turbines in Column 3 by including a dummy variable = 1 if a property lies within 1 mile of a post-construction wind turbine. Column 4 drops all transactions occurring before 2009, and Column 5 excludes all properties that transact more than 4 years before or after the treatment date of the nearest solar installation. All specifications include property, month-year, and county-year fixed effects. The number of observations for the difference-in-differences models are: 106,552 in Column 1, 107,924 in Column 2, 107,291 in Column 3, 61,121 in Column 4, and 33,030 in Column 5. The number of observations for the Ever-treated models are: 10,965 in Column 1, 11,526 in Column 2, 11,292 in Column 3, 6,427 in Column 4, and 3,506 in Column 5. Standard errors are clustered at the tract level and shown in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Appendix

This appendix provides supplemental equations, figures, and tables to our main results.

Figure A1 depicts the increase in new and cumulative solar capacity over time by prior land use.

Figure A2 recreates Figure 3 from our main manuscript, using the same DID specification, but with control defined as 1 – 2 miles. Our conclusion is the same: that the treatment effect does not persist beyond 0.6 miles.

Figure A3 presents a version of Figure 4 from the main manuscript, except with treatment bins defined by 1 year intervals. Our conclusion is the same, treatment begins six months prior to the solar installation operation date.

Figure A4 shows the location of utility-scale solar and wind installations in Massachusetts and Rhode Island.

Figure A5 illustrates the distribution of capacity sizes (in MW) of solar arrays in our data.

Table A1 checks how sensitive our main results are with respect to the definition of the never-treated control group. Each column represents different spatial boundaries of the control group. We find that the treatment effect is significant across the board and very close in magnitude to our main result, implying that our results are robust to different definitions of the control group.

Table A2 explores how different population density cutoff values that define the variable $Rural$ affect the results presented in Panel B of Table 3 in the main paper. Panel A presents results using the sample that includes never-treated observations while Panel B excludes them. 850 people/square mile is the cutoff used in the main text. Overall, the results communicate the same story as the results presented in Table 3: negative impacts are more pronounced in rural areas. All coefficients on $Treated \times Post \times Rural$ are negative and statistically significant. All coefficients on $Treated \times Post \times Non-rural$ are negative, statistically insignificant, and smaller in magnitude than their corresponding $Treated \times Post \times Rural$ coefficient. Magnitudes of the $Treated \times Post \times Rural$ coefficients decline as the population density cutoff increases, which makes sense because more suburban areas are being classified as rural and these areas have smaller impacts.

Table A3 further explores heterogeneity by prior land use and rural character. We estimate a quadruple difference model that interacts the treatment effect term in Equation 1 with the $Non-greenfield$, $Greenfield$, $Rural$, and $Non-rural$ indicator variables. The coefficient on $Treated \times Post \times Non-rural \times Non-greenfield$, which represents the effect of non-greenfield solar arrays in non-rural areas, is negative throughout, though the coefficients are typically small in magnitude and are never significant. The coefficient on $Treated \times Post \times Non-rural \times Greenfield$, which applies to greenfield sites in non-rural areas, is negative across the board but significant only in Columns 1 and 2 at the 5% and the 10% level, respectively. The coefficient on $Treated \times Post \times Rural \times Non-greenfield$, which applies to non-greenfield sites in rural areas, is negative across all specifications with a large magnitude everywhere except Column 4. It is also only significant (at the 10% level) in Column 3. Lastly, the coefficient on $Treated \times Post \times Rural \times Greenfield$, which applies to greenfield sites in rural areas, is statistically significant across the board. This coefficient ranges between -2.1% and -2.5% in Columns 1 – 3, and between -6.4% to -7.5% in Columns 4 and 5. The large coefficients in the sample excluding the never-treated group are consistent with our main results. Overall, these results suggest that the average results are driven by greenfield and rural developments, but no clear picture emerges that one of these attributes is dominant. Instead, there appears to be an additive effect.

Table A4 explores heterogeneity in treatment effect by the size of the solar installations. We define $SmallCapacity$ as an indicator variable = 1 if the size of the installation (in MW) is less than or equal to the median value in our sample (2 MW) and $LargeCapacity$ as an indicator variable = 1 if the size of the installation is greater than 2 MW. We find that the difference between the two categories is small and is statistically insignificant, implying no additional disamenities from solar developments larger than 2 MW. We additionally explore an alternative specification (results not provided) where capacity is treated as a linear variable and is interacted with $Treated \times Post$. These estimates yield the same conclusion to those in Table A3. This result indicates that the presence of utility-scale solar is a disamenity regardless of size. Given that the smallest installations in our analysis are still quite large at five acres in size (about 3.8 football fields), it could be that there is no additional impact of size because it is difficult or even impossible to see beyond five acres from ground level. However, one limitation of this analysis is that the range of observed sizes is narrow. Of the 282 installations in our dataset, almost 47% have a capacity of 2 MW or lesser, and only 23 (8%) are 5 MW or larger. See Figure A6 for a full sense of the distribution of array sizes.

Table A5 examines heterogeneity in treatment effect by time elapsed since treatment. We split our $Post$ variable into two sub-categories: $Post\ (Less\ than\ 3\ years)$ and $Post\ (3\ or\ more\ years)$, where $Post\ (Less\ than\ 3\ years)$ is a dummy variable = 1 if a property transacts less than three years after the treatment date, and $Post\ (3\ or\ more\ years)$ is a dummy variable = 1 if a property transacts 3 or more years after the treatment date. We interact both variables with $Treated,$ and find that the coefficient on $Treated \times Post\ (3\ or\ more\ years)$ is almost similar in magnitude to $Treated \times Post\ (Less\ than\ 3\ years)$ but not always significant, and never are the coefficients statistically significantly different from each other. These results suggest that the treatment effect does not weaken with time.

Table A6 examines anticipation effects leading up to the actual construction of solar installations. We make an ad hoc assumption that the siting process for a solar development can take up to two years. We account for anticipation effects by including an additional term $Treated \times Anticipation$ that interacts $Treated$ with the indicator variable $Anticipation$ which is equal to 1 if a house sells 24 to 6 months prior to solar installation operation date. We find that the coefficient on $Treated \times Anticipation$ is small in magnitude and insignificant across all specifications. Our original treatment effect term $Treated \times Post$ remains mostly unaffected, implying the absence of any prior anticipation effects.

Table A7 checks the robustness of our main results to controlling for differences in price trends between rural and more urban locations. We classify Boston, Providence, Springfield, and Worcester as major cities and include an additional term which is an interaction between distance to nearest city and a year trend in all specifications. This new variable does show up as statistically significant in three of five specification, but our treatment effect coefficients are quite similar to the main results, implying that our results are robust to this type of trend in prices.

Figure A1: New and cumulative capacity by year and land use

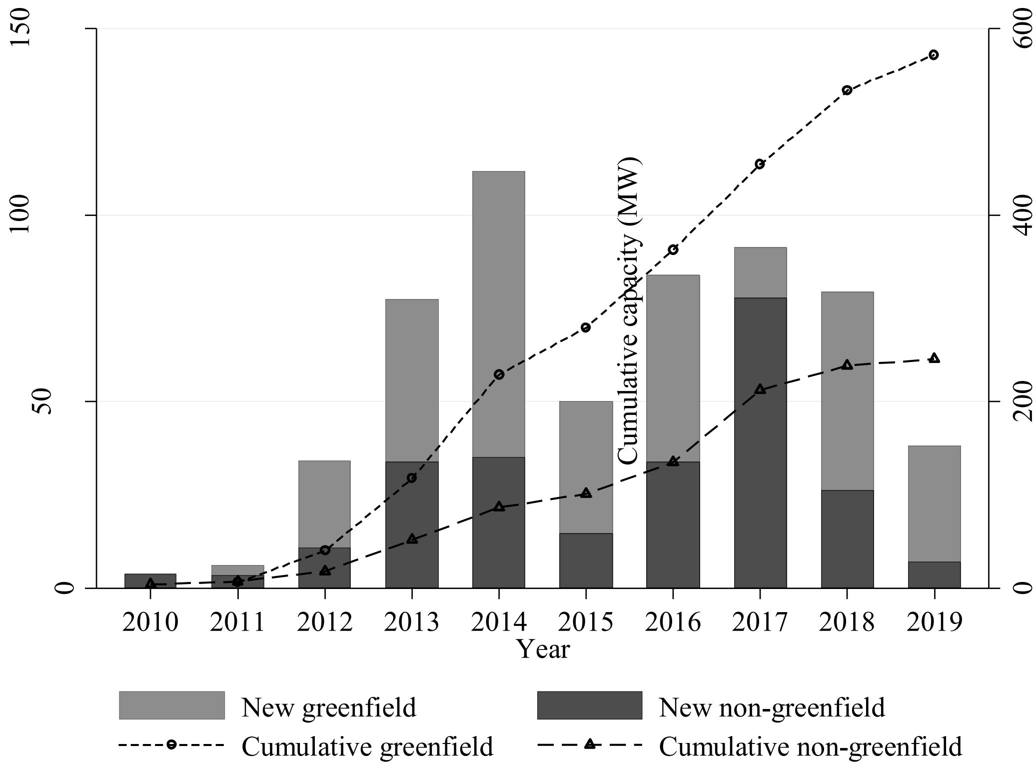
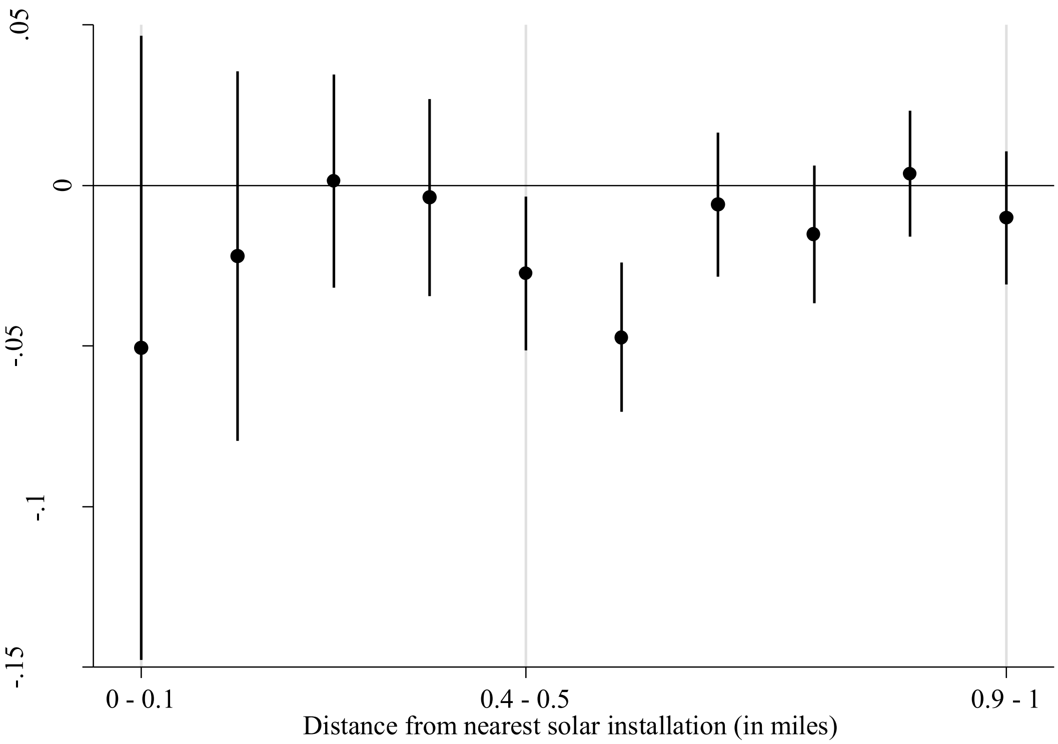
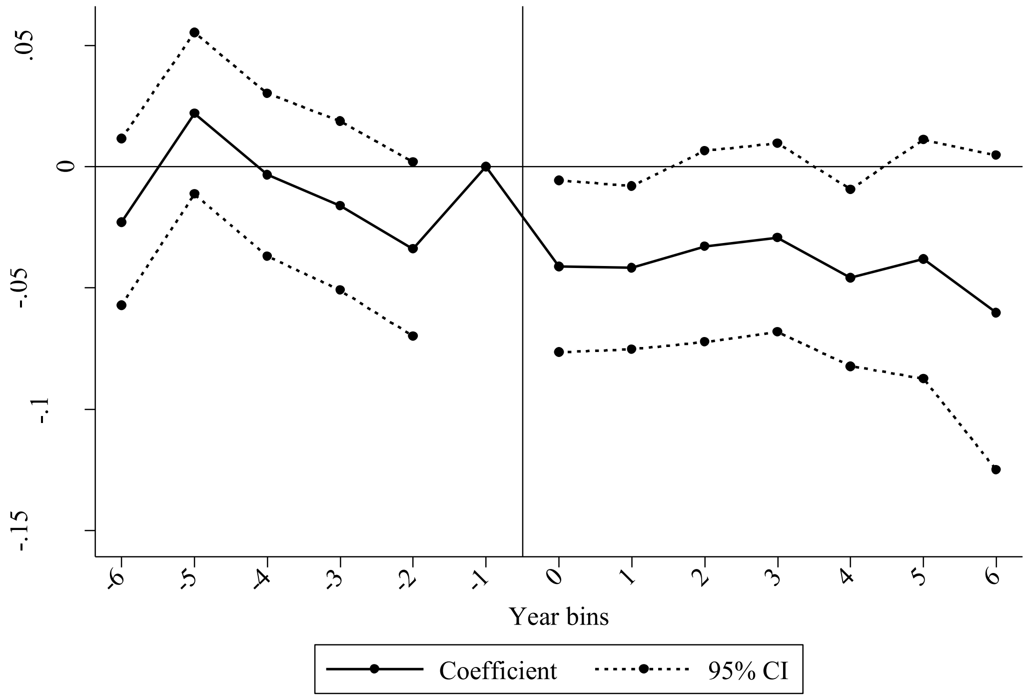


Figure A2: Spatial extent of treatment effect



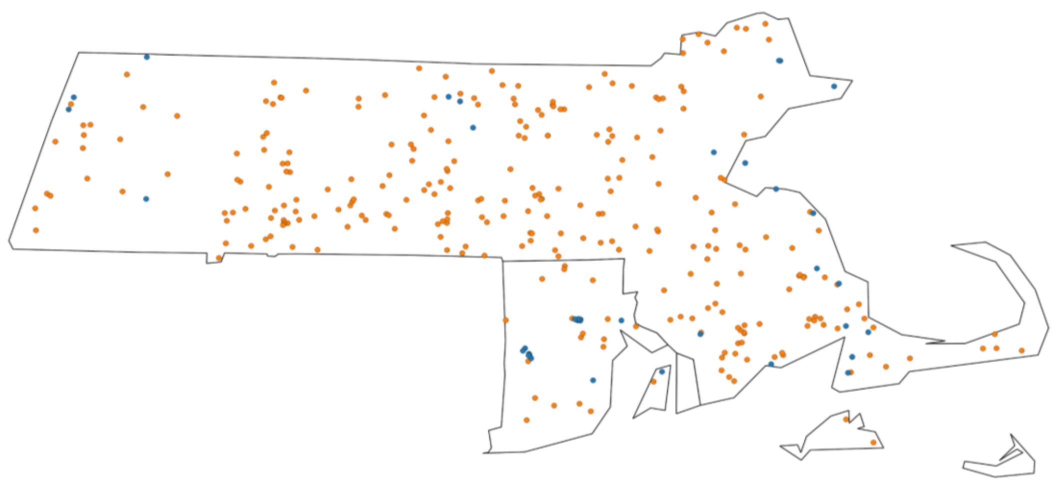
Notes: The treatment variable is defined as a bin variable, with treated properties lying within 1/10 mile distance bands up to 1 mile. Control properties are those lying 1 – 2 miles away from the nearest solar installation. Post = 1 if a house sells after the treatment date. The treatment date is defined as 6 months prior to solar installation operation date. The coefficients are obtained by estimating a DID model similar to Equation 1 that regresses log sales price on 1/10 mile distance bands up to 1 mile, along with month-year and property fixed effects. Resulting coefficients and 95% confidence intervals are graphed.

Figure A3: Event study of prices before and after solar installation operation date



Notes: The treatment variable is defined as a dummy variable equal to 1 if a house is within 0.6 miles of an eventual solar installation site. The time period variable is defined as a bin variable, starting from 6 years prior to solar installation operation date and up to 6 years post operation. Properties are sorted into the respective 1-year bin in which they transact, and the reference time period is 0.5 to 1.5 years prior to operation date. The coefficients are obtained by estimating a DID model similar to Equation 1 that regresses log sales price on the interaction between the treatment and the time period variables, along with month-year and property fixed effects. Resulting coefficients and 95% confidence intervals are graphed.

Figure A4: Location of utility-scale solar and wind installations in MA and RI



Notes: Orange dots represent the location of solar installations and blue dots represent wind installations. Data come from EIA.

Figure A5: Frequency of solar installations by capacity

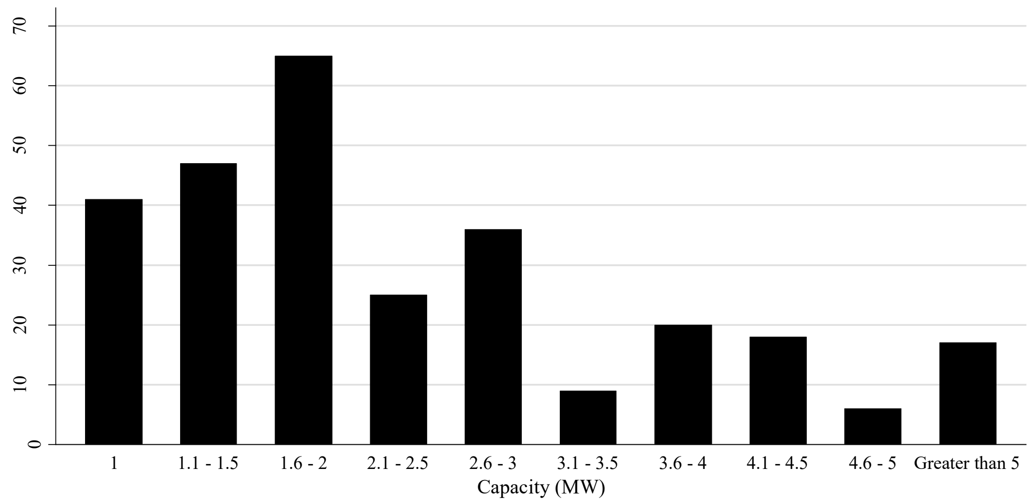


Table A1: Difference-in-differences robustness checks with varying control group bin definitions

Independent variables	Dependent variable: Sale price (ln)			
	Control group distance bands			
	0.6 - 1.5 miles	0.6 - 3 miles	1 - 2 miles	1 - 3 miles
Treated × Post	-0.015**	-0.017**	-0.015**	-0.016**
	(0.007)	(0.007)	(0.007)	(0.007)
Observations	67,836	183,566	86,438	162,713
R^2	0.879	0.875	0.877	0.875

Notes: Treated = 1 if a house is within 0.6 miles of a solar construction, and Post = 1 if a house sells after the treatment date. The treatment date is defined as 6 months prior to solar installation operation date. All models include month-year, county-year, and property fixed effects. Standard errors, clustered at the tract level, are in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Table A2 Heterogeneity of treatment effects by population density

Independent variables	Population density per square mile cutoff					
	500	850	1000	1200	1500	2000
Panel A: Include never-treated						
Treated × Post × Non-rural	-0.010	-0.006	-0.005	-0.007	-0.012	-0.010
	(0.008)	(0.009)	(0.009)	(0.010)	(0.012)	(0.013)
Treated × Post × Rural	-0.031**	-0.025**	-0.025**	-0.022**	-0.016*	-0.017**
	(0.016)	(0.010)	(0.010)	(0.009)	(0.008)	(0.008)
Panel B: Exclude never-treated						
Treated × Post × Non-rural	-0.013	-0.005	-0.008	-0.008	-0.008	-0.003
	(0.013)	(0.015)	(0.015)	(0.015)	(0.017)	(0.018)
Treated × Post × Rural	-0.084***	-0.058***	-0.050***	-0.048***	-0.040***	-0.038***
	(0.019)	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)

Notes: Dependent variable is log sale price and the number of observations is 107,291 in all specifications in Panel A and 11,292 in all specifications in Panel B. Treated = 1 if a house is within 0.6 miles of a solar construction Post = 1 if a house sells after the treatment date. The treatment date is defined as 6 months prior to solar installation operation date. Rural = 1 if the population density per square mile is ≤ column heading value and Non-rural = (1-Rural). All models include month-year, county-year, and property fixed effects. In Panel A, the percentage of observations that qualify as rural for each cutoff value of density per square mile are: 16% for 500, 35% for 850, 39% for 1000, 43% for 1200, 54% for 1500, and 61% for 2000. For Panel B, these percentages are: 20% for 500, 40% for 850, 44% for 1000, 47% for 1200, 58% for 1500, and 65% for 2000. Standard errors, clustered at the tract level, are in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Table A3: Heterogeneity of treatment effects by land use and population density

Independent variables	Dependent variable: Sale price (ln)				
	Include never-treated			Exclude never-treated	
	(1)	(2)	(3)	(4)	(5)
Treated × Post × Non-rural × Non-greenfield	-0.012	-0.006	-0.005	-0.020	-0.003
	(0.016)	(0.010)	(0.011)	(0.018)	(0.017)
Treated × Post × Non-rural × Greenfield	-0.035**	-0.027*	-0.007	-0.033	-0.009
	(0.016)	(0.016)	(0.013)	(0.023)	(0.017)
Treated × Post × Rural × Non-greenfield	-0.026	-0.026	-0.034*	-0.006	-0.041
	(0.020)	(0.019)	(0.019)	(0.038)	(0.025)
Treated × Post × Rural × Greenfield	-0.023*	-0.025**	-0.021*	-0.075***	-0.064***
	(0.012)	(0.012)	(0.012)	(0.021)	(0.015)
Controls					
Month-year fixed effects	Y	Y	Y	Y	Y
Property fixed effects	Y	Y	Y	Y	Y
Municipality time trends		Y		Y	
County-year fixed effects			Y		Y
Observations	107,291	107,291	107,291	11,292	11,292
R^2	0.871	0.876	0.878	0.889	0.891

Notes: Treated = 1 if a house is within 1 mile of a solar construction, and Post = 1 if a house sells after the treatment date. The treatment date is defined as 6 months prior to solar installation operation date. Columns 1 – 3 include the following additional interactions: Treated*Rural, Treated*Greenfield, Post*Rural, Post*Greenfield, Rural*Greenfield, Post*Greenfield*Rural, and Treated*Rural*Greenfield. Standard errors, clustered at the tract level, are in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Table A4: Heterogeneity of treatment effects by solar installation size

Independent variables	Dependent variable: Sale price (ln)				
	Include never-treated			Exclude never-treated	
	(1)	(2)	(3)	(4)	(5)
Treated × Post × SmallCapacity	-0.018	-0.018**	-0.017*	-0.035**	-0.038***
	(0.012)	(0.008)	(0.010)	(0.016)	(0.014)
Treated × Post × LargeCapacity	-0.031***	-0.018*	-0.013	-0.037**	-0.018
	(0.011)	(0.010)	(0.010)	(0.016)	(0.015)
Controls					
Month-year fixed effects	Y	Y	Y	Y	Y
Property fixed effects	Y	Y	Y	Y	Y
Municipality time trends		Y		Y	
County-year fixed effects			Y		Y
Observations	107,291	107,291	107,291	11,292	11,292
R^2	0.871	0.876	0.878	0.889	0.891

Notes: Treated = 1 if a house is within 0.6 miles of a solar construction, and Post = 1 if a house sells after the treatment date. The treatment date is defined as 6 months prior to solar installation operation date. SmallCapacity = 1 if the capacity of the installation is ≤ 2 MW and LargeCapacity = (1- SmallCapacity). Standard errors, clustered at the tract level, are in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Table A5: Heterogeneity of treatment effects by years since construction

Independent variables	Dependent variable: Sale price (ln)				
	Include never-treated			Exclude never-treated	
	(1)	(2)	(3)	(4)	(5)
Treated × Post (Less than 3 years)	-0.027***	-0.020**	-0.018**	-0.037***	-0.028**
	(0.010)	(0.008)	(0.009)	(0.013)	(0.013)
Treated × Post (3 or more years)	-0.010	-0.022*	-0.016*	-0.042**	-0.021
	(0.009)	(0.011)	(0.010)	(0.021)	(0.017)
Controls					
Month-year fixed effects	Y	Y	Y	Y	Y
Property fixed effects	Y	Y	Y	Y	Y
Municipality time trends		Y		Y	
County-year fixed effects			Y		Y
Observations	107,291	107,291	107,291	11,292	11,292
R^2	0.871	0.876	0.878	0.889	0.891

Notes: Treated = 1 if a house is within 0.6 miles of a solar construction. Post (Less than 3 years) = 1 if a house sells within 3 years post-treatment date, and Post (3 or more years) = 1 if a house sells 3 or more years after the treatment date. The treatment date is defined as 6 months prior to solar installation operation date. Standard errors, clustered at the tract level, are in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Table A6: Estimates of the impact of solar installations on property prices including anticipation effects

Independent variables	Dependent variable: Sale price (ln)				
	Include never-treated			Exclude never-treated	
	(1)	(2)	(3)	(4)	(5)
Treated × Anticipation	-0.003	0.006	0.003	0.004	-0.005
	(0.010)	(0.011)	(0.010)	(0.017)	(0.014)
Treated × Post	-0.024***	-0.017**	-0.014**	-0.033*	-0.032*
	(0.009)	(0.007)	(0.007)	(0.018)	(0.017)
Controls					
Month-year fixed effects	Y	Y	Y	Y	Y
Property fixed effects	Y	Y	Y	Y	Y
Municipality time trends			Y		Y
County-year fixed effects		Y		Y	
Observations	107,291	107,291	107,291	11,292	11,292
R^2	0.871	0.876	0.878	0.889	0.891

Notes: Treated = 1 if a house is within 0.6 miles of an eventual solar installation site, Post = 1 if a house sells after the treatment date, and Anticipation = 1 if a house sells 6 to 24 months prior to treatment date. The treatment date is defined as 6 months prior to solar installation operation date. Standard errors are clustered at the tract level and shown in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Table A7: Estimates of the impact of solar installations on property prices accounting for distance to city by year trends

Independent variables	Dependent variable: Sale price (ln)				
	Include never-treated			Exclude never-treated	
	(1)	(2)	(3)	(4)	(5)
Treated × Post	-0.024***	-0.018***	-0.013**	-0.036***	-0.028**
	(0.008)	(0.007)	(0.007)	(0.013)	(0.013)
Miles to city × year	0.0001**	-0.0002	-0.0004***	-0.0001	-0.0004***
	(0.00004)	(0.0002)	(0.0001)	(0.0005)	(0.0001)
Controls					
Month-year fixed effects	Y	Y	Y	Y	Y
Property fixed effects	Y	Y	Y	Y	Y
Municipality time trends		Y		Y	
County-year fixed effects			Y		Y
Observations	107,291	107,291	107,291	11,292	11,292
R^2	0.871	0.876	0.878	0.889	0.891

Notes: Treated = 1 if a house is within 1 mile of a solar construction, and Post = 1 if a house sells after the treatment date. The treatment date is defined as 6 months prior to solar installation operation date. Standard errors, clustered at the tract level, are in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Manuscript – 2

Submitted to Resource and Energy Economics, May 2021

When Energy Issues are Land Use Issues: Estimating Preferences for Utility-Scale Solar Energy Siting

Vasundhara Gaur [a], Corey Lang [a], Gregory Howard [b], and Ruth Quainoo [b]
[a] Department of Environmental and Natural Resource Economics, University of Rhode Island
[b] Department of Economics, East Carolina University

ABSTRACT
Solar energy has grown rapidly in recent years and will continue to do so as states and nations seek to curtail carbon emissions. While solar energy receives broad support in general, the siting of utility-scale solar arrays is contentious because at that stage it becomes a land use issue replete with potential disamenities and difficult tradeoffs. We developed and conducted a choice experiment survey to estimate preferences for various attributes of utility-scale solar arrays in Rhode Island, USA, where ambitious renewable energy targets collide with scarce land resources. Our results suggest the largest indicator of solar development approval is prior land use, with residents willing to pay an additional $10-21 in monthly utility bills for developments in commercial, industrial, brownfield, and covered landfill areas, whereas they are willing to pay $13-49 to avoid developments on farm and forest land. Additionally, respondents are willing to pay about $6-8 per month for a solar array to be fully screened and not visible. We conclude with a discussion of how these preferences can be incorporated into state and local solar siting policy.

Keywords: Solar energy; Utility-scale solar; Willingness to pay; Contingent valuation; Choice experiment
JEL codes: Q24; Q42; Q51

1 Introduction

Solar energy has grown immensely in the United States, with an average annual growth rate of almost 42% since 2010 (M. Davis et al., 2021). In 2020, the United States had over 89 gigawatts (GW) of total installed capacity, which is enough to power 16.4 million homes and accounts for 2.6% of total electricity generation (M. Davis et al., 2021). In the next few years, solar energy is projected to grow faster than any other renewable source in the United States, more than doubling its total installed capacity by 2025, and accounting for 47% of total renewable electricity generation by 2050 (*EIA Annual Energy Outlook*, 2021).

Despite broad support for solar energy in the United States (Carlisle et al., 2014, 2015; Farhar, 1994; Greenberg, 2009; Jacobe, 2013; Pew Research Center, 2019), the construction of utility-scale solar installations (sized 1 MW and above) is often fraught with hurdles. The key insight as to why solar can be divisive is to understand that siting of utility-scale solar is a land use issue as much as it is an energy issue. Utility-scale solar installations require large amounts of land. The proliferation of solar has become the largest cause of land use change in the United States (Trainor et al., 2016). On average, a solar installation with a capacity of one megawatt (MW) requires five acres of land, which is over ten times the land area required by conventional sources (Denholm and Margolis, 2008; Ong et al., 2013) and is often the greatest obstacle to additional solar development. Other concerns of residents related to solar development include glare from glass panels, ecosystem impacts, loss of scenic beauty and rural character, water pollution, and reduction in property values (Abashidze, 2019; Dröes and Koster, 2021; Farhar et al., 2010; Gaur & Lang, 2020;

Gross, 2020; Lovich & Ennen, 2011; Qi & Zhang, 2017; Tsoutsos et al., 2005; Turney & Fthenakis, 2011).

The debate regarding utility-scale solar siting is particularly contentious in Rhode Island (RI), which is the setting of our study. In 2004, RI adopted an ambitious Renewable Energy Standard, which set the goal of generating 38.5% of total energy from renewable sources by 2035. To this end, 80 MWs of utility-scale solar energy capacity have been built since 2013 (EIA, 2021), and the pace of development has increased recently (Kuffner, 2018). Much of the concern regarding solar energy expansion stems from the fact that the most common sites for arrays are on forest and farmlands. While these are the areas where development is cheapest, they offer many amenities to residents, particularly in a small state with scarce land resources, the nation's second highest population density, and strong public support for land conservation and environmental preservation (Altonji et al., 2016).

The purpose of this paper is to quantify the perceived externalities from utility-scale solar installations by estimating the tradeoffs people are willing to make for a set of siting attributes. We designed a choice experiment (CE) and administered a survey to estimate preferences for four siting attributes of utility-scale solar installations: size of the installation, visibility, setback distance, and current land use of the proposed development site. The four land types we consider are forest land, farmland, commercial/industrial land, and brownfields/covered landfills, which represent standard siting options. The survey presents respondents with multiple hypothetical solar development plans with different attributes (including a no solar development alternative) and asks their preferred option. Each alternative is paired with a change in household electricity bill, and thus respondents are making tradeoffs between money and solar attributes. Through their choices, we can estimate an average monetary value (willingness to pay, or WTP) for each solar siting attribute.

Our results are consistent with expectations, but the specific magnitudes of willingness to pay yield insights into Rhode Islanders' priorities. The results indicate that respondents prefer larger installations and are willing to pay about $1.25 per month per MW of solar energy capacity, which demonstrates overall support for continued transition to solar energy. Respondents dislike visible installations and are willing to pay between $6.21 and $8.42 per month to avoid a nearby installation that is completely visible. Our results suggest the largest factor in determining approval is current land use of the proposed development site, and there is substantial heterogeneity across land types. Respondents have a preference for solar installations sited on brownfields and commercial lands, with an average WTP between $10.08 to $15.11 for brownfields and $14.48 to $20.78 for commercial areas. In stark contrast, they are willing to pay to avoid solar development on forest lands and farmlands. Conversion of forest land is most detested, and estimated WTP ranges from $40.60 to $49.10 per month.

While these results are consistent with expectations and sentiment expressed in town meetings and from stakeholders, they are important because they quantify resident preferences in a way that can guide statewide policy and local siting ordinances. Specifically, states can offer additional subsidies for solar development on industrial/commercial areas, brownfields, and covered landfills, which are necessary to entice developers because arrays on these sites are more expensive to build and maintain. Importantly, our analysis shows that economically meaningful subsidies are highly likely to pass a benefit-cost analysis under reasonable assumptions and are thus warranted in a social welfare framework. Visual screening is another important component of development proposal and approval, and our results suggest that the significant costs borne to screen an array are also justified by the benefits of residents not seeing the array. We discuss these policy ramifications in Section 6 in more detail.

This study advances the literature in several ways. First, we provide the first estimates of the valuation of utility-scale solar siting attributes in the United States. Even beyond the United States, literature on non-market valuation of utility-scale solar energy is sparse. To date, there are only four studies that use a CE to estimate externalities from utility-scale solar: Botelho et al. (2017) in Portugal, Yang et al. (2017) and Kim et al. (2020) in South Korea, and Oehlmann et al. (2021) in Germany. Botelho et al., (2017) estimate a marginal WTP to avoid glare from solar panels of \$5.15 per month. Yang et al., (2017) also find a negative WTP for light pollution caused by glass arrays, estimating a value of \$14 per household per month. Our finding that respondents need to be compensated between \$6.21 and \$8.42 per month for a completely visible installation falls within the range estimated in previous studies. Oehlmann et al. (2021) find that respondents in Germany prefer solar to wind or biomass. Further, we are the first research in the solar siting literature to translate household WTP estimates, the standard in this literature, into a more actionable unit of measure that reflects subsidies in this area, specifically an aggregate WTP/kWh for an installation.

Our second contribution, and the largest difference between our study and those prior, is our explicit distinction between possible prior land uses. Previous studies all find that solar installations have a negative impact on the landscape, though each defines the "landscape" attribute and its associated levels differently. Botelho et al. (2017) consider a general kind of landscape without differentiating between land use types and find that the average WTP to avoid "significant impacts on landscape" is $7.58, relative to no impact. Similarly, Yang et al. (2017) also assume a general definition of "landscape destruction" and define levels in terms of percentage decreases in natural beauty. They estimate a WTP of $0.05 per percentage point of landscape destruction. Kim et al. (2020) capture landscape impacts on flatlands (farmlands and orchards) and mountainous areas and find that people need to be compensated $1,951 per month for solar development on flatlands, and $1,059 per month for solar on mountainous lands, compared to solar panels located on rooftop and walls. Oehlmann et al. (2021) focus mostly on proximity and do not explicitly describe prior land use. Further, the status quo still involves renewable energy development, just of an uncertain type, which makes it unclear how respondents assess land use impacts. While each study captures landscape impacts differently, none estimate the impact of current land use on preferences for solar development in a way that is theoretically rigorous (understanding that the perceived disamenity of solar development is a function of the lost amenity values of the current land use) and framed in a manner that is equally understandable to residents and actionable for policy makers.

Lastly, our research contributes broadly to our understanding of NIMBYism, which is loosely defined as when people approve of something in general, but do not want it near their residence. Objections to renewable energy siting are frequently characterized (and maybe dismissed) as merely NIMBYism. However, this simplistic view often does not hold up to scrutiny. Devine-Wright, (2009) argues instead that local opposition is about protecting places that people are attached to and derive meaning and identity from. Boyle et al. (2019) conduct a choice experiment about onshore wind energy siting and reject a NIMBY hypothesis because those that have positive views of wind in general are more likely to choose development of wind locally. We add additional nuance to this debate. Our respondents overwhelmingly (87%) favor solar energy in general but are more discerning when it comes to local projects, approving solar in only about 60% of choice tasks. These statistics alone could indicate NIMBY attitudes, but our results tell a nuanced story that hinges on land use. We find that people have positive WTP for local solar on brownfields and commercial areas, but these same people require compensation for solar developed on farm and forest lands. Thus, support is not driven exclusively by proximity, and so does not render itself to a purely NIMBY explanation, but on other development plan characteristics.

2 METHODS

2.1 Choice experiment design

In our CE, we present each respondent with six choice tasks modeled as solar development plans. Each development plan asks respondents to consider a hypothetical group of land parcels that have three main characteristics. First, all land parcels are near each other and total fifty acres. Second, they are less than fifteen minutes from the respondent's residence by car in Rhode Island. And third, each group of parcels has one of the following four different land types: brownfield, commercial, farmland, and privately owned forest. The survey presents two choice tasks for both farmland and forest parcels, and one each for commercial and brownfield land types. We chose to disproportionately ask about farmland and forest land because these are more common siting locations in New England and the most contentious, so we prioritized these for precise estimates.

Our CE design differs from much of the literature by treating current land use differently than most choice attributes. In pilot testing our survey, we received feedback from stakeholders suggesting that a realistic portrayal of solar development decisions would involve different development plans on a single site (and so a single current land use) rather than development plans involving different land uses. Another concern, largely technical in nature, with the traditional CE design was that it was unclear how land use should be coded for the status quo alternative (or the no solar option) when multiple development options with different land uses were presented in a single choice. We tackle this issue by keeping current land use constant between different development plans in a given choice while varying land use between choices. Econometrically, this means we are unable to include land use variables in our choice model as one would include other attribute variables because they lack within-choice variation. However, we are still able to recover the impact of land use differences from our choice model by interacting land use indicator variables with alternative-specific constants (ASCs). As such, our work highlights a method of including in a choice design and estimating preferences for attributes that may not be credibly varied within a given choice set.

We develop a D-efficient design using Stata, which included 30 choice sets for commercial and brownfield land types. The farm and forest land use designs included 60 choice sets, which were divided into blocks of two questions. In pilot testing the survey, both with focus groups and with an advisory group of stakeholders knowledgeable about solar development in Rhode Island, we identified several areas where specific attribute levels did not make sense for certain land uses. In developing our experimental design for each land use type, we specified these constraints, then allowed our software to identify the D-efficient design given these constraints.

For each choice opportunity, we present three hypothetical development plans, labeled A, B, and C. The first two plans assume that some or all of the parcels of land under consideration will be developed into utility-scale solar installations, but with varying solar installation characteristics. The final alternative (Choice C) is a status-quo option where the land will be free of solar panels and will remain in its current use 'for the time being'.

Our CE presents four solar attributes: *size of installation*, *visibility*, *setback*, and *change in electricity bill. Size of installation* indicates the area of land (in acres) that is converted to solar energy production, and how many households are capable of being powered by the installation under consideration. *Visibility* refers to how visible a solar installation is from the respondents' house or from regularly traveled roads. *Setback* is the minimum distance of the solar panels from the property line. The attribute representing our payment vehicle is *change in electricity bill*, which is defined as the dollar increase or decrease in respondents' electricity bill if a specific development plan is implemented. For ease of understanding, we present the change in both monthly and annual terms. Finally, our CE also includes the attribute *probability of residential development* when the land type is either farmland or forest. This is because most farm and forest land is zoned residential in Rhode Island, and there is a possibility that it will be converted into residential housing in the future if it is not developed into solar. This attribute was added based on discussions in focus groups and represents the reality that privately held land may not remain open space indefinitely. Figure 1 shows an example choice set for the farmland land use.[16] Table 1 defines all attributes and their associated levels used in our design.

Our survey is divided into four sections. The first section provides background information about our study and the history of siting utility-scale solar installations in Rhode Island. We convey that the objective of the study is to help policy makers implement decisions that reflect the public's preferences, and that the final results will be disseminated to state and local decision makers and the public at large through outreach.[17] We also inform participants that our study is backed by an advisory group consisting of officials in state and local governments, non-profit environmental organizations, and solar development experts who have also provided guidance at various stages of the project. In the second section we ask respondents questions regarding their energy usage and attitudes about different energy sources. The third section first defines each attribute in our CE and familiarizes respondents with its overall structure using an example choice, then presents the six choice questions in a

[16] Figures A1, A2, and A3 in the online appendix depict example choice sets for the forest, commercial and brownfield land types, respectively. Each choice set was also accompanied by a picture depicting prior land use, but we omit these in the figures out of copyright concern.
[17] The grant that funded this work requires integration of research and extension.

randomized order. The fourth and final section includes questions designed to assess perceived consequentiality of the survey, identify stated attribute non-attendance, and collect demographic information.

2.2 Empirical models

The choices made by respondents in our CE allow us to empirically estimate their preferences for a variety of solar siting attributes. McFadden (1974) laid the theoretical groundwork linking consumers' choices to utility maximization through the Random Utility Model (RUM). In the RUM context, the utility of individuals is assumed to have two components: an observable and an unobservable (random) component. This can be expressed as:

$$U_{im} = V_{im} + \varepsilon_{im} \tag{1}$$

where U_{im} is the utility that respondent i derives from alternative m, which is a function of his observable utility V_{im} and his random utility ε_{im} from choice m. The observable component, V_{im}, can depend on individual-specific characteristics and the attributes of alternative m.

We use the standard multinomial (conditional) logit (CL) model proposed by McFadden (1974) to model respondents' choices. The CL model requires that choices be independent of irrelevant alternatives (IIA) and makes two main assumptions: first, that all individuals have homogenous preferences, and second, that the variance of the error term is constant across individuals. In this case, that probability of individual i choosing alternative m is given by:

$$P_{im} = \frac{exp(\lambda V_{im})}{\sum_{n=1}^{N} exp(\lambda V_{in})} \tag{2}$$

where λ is a positive scale factor that is inversely proportional to the error variance, σ_{ε}^{2}:

$$\lambda = \frac{\pi}{\sqrt{6\sigma_{\varepsilon}^{2}}} \tag{3}$$

When error terms are IID, the error variance, and thus λ, are constant across individuals. Since the scale parameter cannot be directly estimated, it is typically normalized to unity, an assumption that has been called into question in the literature several times (DeShazo and Fermo, 2002; Hensher et al., 1998; J. Louviere et al., 2002; J. J. Louviere, 2001).

To allow error variances (and scale parameters) to vary across individuals and choices, we employ an alternative model known as the heteroskedastic conditional logit (HCL) (DeShazo and Fermo, 2002; Hensher et al., 1998). In this model, scale parameters are represented as:

$$\lambda_{im} = exp(\boldsymbol{\varphi Z_{im}}) \tag{4}$$

where $\boldsymbol{Z_{im}}$ is a vector of individual- and choice-specific characteristics (specified as the four different land use types in our model) and φ is the parameter that describes the effect of those characteristics on the scale parameter. The probability of individual i choosing alternative m then becomes:

$$P_{im} = \frac{exp(\lambda_{im}\boldsymbol{V_{im}})}{\sum_{n=1}^{M} exp(\lambda_{im}\boldsymbol{V_{in}})} \tag{5}$$

Finally, we use the random parameters logit (RPL), or mixed logit model, which relaxes the IIA restrictions of the CL model and additionally allows for preference heterogeneity. It does this by incorporating a random parameter into the utility function that represents how much each individual's preferences deviates from the population mean. Therefore, the utility each individual i gets from alternative m in situation t can be represented as:

$$U_{imt} = \boldsymbol{X_{imt}}(\boldsymbol{\beta} + \boldsymbol{\eta_i}) + \varepsilon_{im} = \boldsymbol{X_{imt}}(\boldsymbol{\beta_i}) + \varepsilon_{im} \tag{6}$$

where $\boldsymbol{X_{imt}}$ represents the observed attributes, $\boldsymbol{\beta}$ is a vector of mean coefficient values associated with those attributes, and $\boldsymbol{\eta_i}$ is a vector of individual-specific deviation parameters that captures preference heterogeneity. Preference heterogeneity is therefore captured directly in the RPL model through the vector $\boldsymbol{\beta_i}$, which represents individual-specific preference parameters for each attribute with assumed preference heterogeneity. The probability of individual $i's$ sequence of choices $[c_1, c_2, \dots . c_T]$ is given by:

$$P_{i[c_1,c_2,\dots c_T]} = \int \dots \int \prod_t^T \left[\frac{exp(\boldsymbol{X_{imt}\beta_i})}{\sum_{n=1}^{M} exp(\boldsymbol{X_{int}\beta_i})}\right] f(\boldsymbol{\beta})d\boldsymbol{\beta} \tag{7}$$

2.3 Estimation

Our main expected utility specification is given as:

$$\begin{aligned} V_{im} = {} & \beta_{Acres}Acres_{im} + \beta_{PartVis}PartVisibility_{im} + \beta_{FullVis}FullVisibility_{im} \\ & + \beta_{Setback}Setback_{im} + \beta_{Probability}Probability_{im} + \beta_{Cost}Cost_{im} \\ & + \beta_{FarmASC}Farm_i \times ASC_{im} + \beta_{ForestASC}Forest_i \times ASC_{im} \\ & + \beta_{BrownfieldASC}Brownfield_i \times ASC_{im} \\ & + \beta_{CommercialASC}Commercial_i \times ASC_{im} \end{aligned} \tag{8}$$

where $Acres_{im}$ refers to the size of the installation (in acres), $PartVisibility_{im}$ and $FullVisibility_{im}$ are indicator variables equal to 1 if the installation is partly visible and completely visible,[18] respectively, $Setback_{im}$ refers to the setback distance (in feet), $Probability_{im}$ is the probability of residential development on farm and forest land in lieu of solar development, $Cost_{im}$ is the change in respondents' monthly electricity bill, and ASC_{im} is the status-quo alternative-specific constant, or a dummy variable equal to 1 for the status-quo choice and equal to 0 for either of the solar development options (Choices A and B). $Farm_i$, $Forest_i$, $Brownfield_i$, and $Commercial_i$ are all dummy variables equal to 1 if the choice set is framed around the respective land use.

In Equation (8), each solar attribute k is associated with a preference coefficient β_k, which are estimated using maximum likelihood procedures. The interaction coefficients allow us to identify whether respondents have different preferences (and different WTP's) for each land type. The ASC_{im} term indicates respondent i's desire to choose the status-quo alternative over other solar development alternatives, which can also be interpreted as their dislike for solar arrays. The interaction of the ASC_{im} term with a land use type l will therefore represent their preferences for developing solar arrays on that particular land use type. If the coefficient associated with the interaction between land use type l and the ASC_{im} term (β_{lASC}) is positive, it implies that people prefer the status-quo option over the other alternatives, or equivalently that they dislike having solar arrays on the associated land parcel, all else equal.

These coefficient estimates can be used to make welfare calculations. We obtain the marginal WTP (MWTP) value for a particular attribute k by dividing the

[18] The omitted category is not visible at all.

coefficient of that attribute with the negative of β_{Cost}, the coefficient associated with the cost variable. Mathematically:

$$MWTP_k = -\frac{\beta_k}{\beta_{Cost}} \tag{9}$$

We can also estimate the maximum WTP (also called compensating variation or CV) for a specified plan by finding the price that makes the utility derived from that plan, denoted as V_i^1, equal to the utility from the status quo option, denoted $V_i^{SQ\_l}$. Note that status quo utility is indexed by land use, as our interactions of the SQ ASC with land use allows us to estimate different status-quo utilities for different land uses. From here we can estimate a unique CV for each land use type l:

$$CV_l = \frac{V^{SQ\_l} - V^{1*}}{\beta_{Cost}} \tag{10}$$

where V^{1*} is the utility of the non-price attributes associated with the solar development plan under consideration. Subtracting the CV associated with one land use type from another gives us the premium the average respondent would pay for switching solar panels from one current land use to the other.

3 DATA

3.1 Survey implementation

We use the Tailored Design Method formulated by Dillman et al. (2014) to design a mixed-mode, web-push survey. The mixed mode aspect enables us to collect data both online (using Qualtrics) and through mail, allowing for a higher response rate and greater sample representativeness (Millar and Dillman, 2011). The web-push aspect allows us to contact potential respondents by mail and invite them to take the survey online, which lowers per-respondent cost (McMaster et al., 2017).

We drew a random sample of 3,000 individuals from the 2019 Rhode Island voter registration database, which is publicly available from the Secretary of State. These data include name, address, age, party affiliation, and whether the individual participated in the last eight elections held. Sample selection probabilities were adjusted to increase the odds of selecting younger people, those living in rural areas, and Republicans. Republicans were oversampled because they are a smaller group in Rhode Island and have been found to be less responsive to surveys (Best and Krueger, 2012; Pearson-Merkowitz and Lang, 2020). Rural residents were oversampled because they are more likely to be impacted by solar siting decisions. We oversampled younger residents because we anticipated lower response rates from them.

We disseminated the survey in three rounds. The first round was mailed on September 4, 2020. Each envelope included an introductory letter that provided a link and unique access code to the online survey and a $2 cash incentive. Two weeks later, non-respondents were sent a follow-up postcard as a reminder, which also gave the link and access code. In the third and final round (mailed two weeks after the second round), subjects who had not responded to either of the first two rounds of mailings were sent a paper survey along with a prepaid, pre-addressed return envelope.

Of the 3,000 surveys that were mailed, 204 were returned as non-deliverable. We received 669 total responses (24% response rate), 510 of which came from the online mode and 159 from mail. We drop 13 individuals who do not answer any of the choice questions. Our final sample consists of 3,914 choices made by 656 individuals.[19]

3.2 Summary statistics

Summary statistics of respondent characteristics are presented in Table 2. The average annual household income is $109,250 and the average monthly electricity bill is $123.57. About 68% of the respondents have a college degree or higher, 63% are employed, and 52% are female. A large proportion of respondents are homeowners (83%), 35% have children at home, and the average tenure in their current home is over 15 years. About 5% of subjects have solar panels installed in their own homes. On average, they have a very positive attitude towards renewable energy sources (solar, wind, and hydro). Specifically, 87% of respondents have a positive view of solar energy in general. In contrast, respondents are neutral towards natural gas and dislike energy production from nuclear materials and coal. These attitudes are consistent with recent nationwide studies that find immense support for developing alternative energy over expanding fossil fuels in the U.S (McDonald et al., 2020; Pew Research Center, 2020). Finally, over 90% of subjects find the survey to be consequential with regards to policy decisions and the payment vehicle.[20]

To make our sample representative of the Rhode Island population, we use the voter registration data to construct sample weights. Three key demographic variables are used to construct weights: age, political affiliation, and rural/urban residence.

[19] Only 36 respondents chose the status quo alternative in all six choice questions, giving us a serial non-participation rate of 5.5%, which is considerably lower than other studies (Chen et al., 2020; von Haefen et al., 2005). This finding suggests that respondents are engaging with the subject and not dismissing it outright. Figure A4 in the online appendix depicts respondents' choice preferences for the status quo and solar development alternatives by land use.

[20] Following Carson and Groves (2007) and Herriges et al. (2010), we use a knife-edge definition of consequentiality where policy consequentiality is an indicator variable equal to 1 if respondents believe that their answers will influence policies. Likewise, payment consequentiality is an indicator variable equal to 1 if respondents believe that they will have to pay with any positive probability.

Table 3 reports the demographic distribution for these three variables in our unweighted sample, the population, and the weighted sample. The unweighted sample means differ from the population means across all groups, which is due to our disproportionate sampling and various groups' propensity to respond to the survey. However, the application of survey weights balances the proportions exactly.

4 RESULTS

Table 4 reports the estimation results for our main specification. In Column 1 we present coefficients from the CL model. Column 2 shows coefficients derived from estimating the HCL model, along with scale parameter estimates associated with farm, forest, and commercial land use types.[21] Columns 3 and 4 report mean coefficient and standard deviation estimates, respectively, from the RPL model, which is our preferred specification because of its more realistic assumptions regarding preference heterogeneity. Results are consistent across columns. We find that the coefficient on $Acres$ is positive and significant (at 1%), implying that respondents prefer large solar installations. They also dislike installations that are visible, as suggested by the negative sign on $PartVisibility$ and $FullVisibility$. However, only the coefficient on $FullVisibility$ is significant (at 1%), indicating that completely visible installations elicit a stronger negative reaction than partly visible ones. The coefficient on $Setback$ is positive across the board, insignificant in the CL and HCL models, and only weakly significant (at the 10% level) in the RPL model. This suggests that people are

[21] Coefficients for land use in "Heteroskedastic variables" portion of this model are read as the change in scale parameter (or, more specifically, the change in the exponent of the scale parameter) for the land use relative to the omitted land use, which is brownfield.

unaffected by setback distance when controlling for the visibility of the installation. This is also likely because respondents consider setback distance to be the least important attribute while making choices (Figure A5 in the online appendix). $Probability$ is negative and significant at the 1% level, which means that people are less likely to choose an option when the probability of residential development is higher. Since the only options in our design with nonzero probability of residential development are status-quo options when forests or farmlands are the current land use, the implication is that respondents are less likely to select the status-quo (and so more interested in solar development) if the land is more likely to be converted to housing in the near future, which is consistent with expectations.

We find that preferences for constructing solar installations differ by the type of land use under consideration. The positive and significant coefficient on $Farm \times ASC$ in the CL and HCL models suggests that respondents' prefer the status-quo for this land use type, and thus dislike having solar arrays built on farmlands. The corresponding mean estimate for the RPL model is positive and significant only at the 10% confidence level, though the large and significant SD value implies that people exhibit large variation in their preferences regarding solar installations on farmlands. The coefficient on $Forest \times ASC$ is positive and highly significant across all models, providing strong evidence of people's dislike for developing forest lands for solar energy. Similar to farmlands, we find evidence of large variation in respondents' preferences for converting forest land into solar installations, as indicated by the large and significant SD values associated with the $Forest \times ASC$ term. The negative and significant (at the 1% level) coefficients on the $Brownfield \times ASC$ and

$Commercial \times ASC$ interaction terms indicate that, in general, people like having solar installations on brownfields and commercial land types.

In Panel A of Table 5 we present MWTP estimates for all attributes with standard errors derived using bootstrapping with 1,000 replications. On average, respondents are willing to pay \$0.24 to \$0.28 per month for each additional acre of land to be developed for solar. This translates to a monthly WTP between \$7.20 and \$8.40 for a 30 acre installation and between \$12 and \$14 for a 50 acre one, which, in a basic sense, is consistent with overall support for solar energy and general subsidies for solar energy. We find that the MWTP for a partly visible installation is negative, though insignificant, and small in magnitude. The MWTP for a fully visible installation is significant and much larger in magnitude, which suggests that respondents need to be compensated between \$6.21 and \$8.43 per month for solar installations that are completely visible, compared with not visible. Values for the $Setback$ attribute are insignificant for the CL and HCL models, and slightly significant (at the 10% level) for the RCL model. However, the magnitude is small throughout, implying that respondents are largely unaffected by setback distance. The MWTP for $Probability$ is negative and significant, indicating less compensation is needed for solar developed on farm and forest lands when the probability of future residential development increases. In addition, these estimates can be interpreted as MWTP for permanent land conservation. On average, respondents are willing to pay between \$4.75 and \$11.25 per month for a 25% reduction in the probability of future residential development, and between \$9.50 and \$22.50 per month for a 50% reduction. Translating these monthly payments in perpetuity into present discounted value yields amounts that are similar to property values studies on the capitalization of conserved open space (Irwin, 2002; Lang, 2018). Our MWTP estimates are also broadly similar to various contingent valuation studies estimating the value of farmland and forest conservation across several countries (Jin et al., 2018; Lehtonen et al., 2003; Shoyama et al., 2013).

Panel B of Table 5 reports CV estimates for the development of solar on various land types. We assume a 10 acre solar installation that is completely visible, has a setback distance of 150 feet and with a 0% probability of residential development in the future. Our results provide suggestive evidence of respondents' dislike for constructing solar panels on farmland. Estimates from the CL and HCL models suggest that people need to be compensated almost \$23 per month when farmland are converted to solar installations. In comparison, the RPL estimate of \$13.22 per month is smaller in magnitude, though it is still negative and significant. We find large negative WTP values for constructing solar on forest lands, which indicates a strong dislike for such siting. On average, people need to be compensated between \$40.58 and \$49.04 per month for the development of forest land into solar. We also find positive WTP values for $Commercial$ and $Brownfield$, implying that respondents support converting these types of lands into solar installations. Our results indicate that people are willing to pay between \$14.43 and \$20.72 per month in higher energy bills for solar development on commercial lands and range from \$10.06 to \$15.07 per month on brownfields.

Given the dominance of land use in determining project approval, we additionally investigate whether attribute preferences vary by land use. We split the sample of choice tasks by land types revealed to be desirable (commercial and brownfield) and undesirable (farmland and forest), and we estimate Equation (8) on each sample separately. Table 6 presents WTP values for solar attributes for the farm and forest subsample in Column 1 and for the commercial and brownfield subsample in Column 2.[22]

We find several differences across columns that reveal how land use impacts MWTP for attributes. While respondents are indifferent about the size of the installation when built on farm and forest lands, they are willing to pay \$0.38 for each additional acre of solar on commercial and brownfield land. The latter translates to a monthly WTP of \$11.40 for a 30 acre installation and \$19 for a 50 acre one. Visibility is more of a concern to respondents for farm and forest sites than on commercial and brownfield sites. Respondents are willing to pay \$3.38 to avoid seeing a partially visible installation on farm and forest lands, whereas the corresponding value for commercial and brownfields is small and statistically insignificant. While completely visible arrays are disliked regardless of the land type on which they are sited, respondents' dislike is stronger for fully visible installations on farms and forest lands. Their monthly WTP to avoid seeing fully visible installations on commercial and brownfield areas is \$4.42 but is \$10.34 for installations on farm and forest lands, which is a ratio of about 2:5. Respondents also prefer greater setback for solar in commercial and brownfield areas, though the result is only marginally significant. MWTP for reduced probability of residential development and CV estimates in Panel B are consistent with our main results.

[22] The conditional logit regression coefficients that are used to create Table 6 are presented in Table A1 of the online appendix.

Purely for comparison purposes, we also develop a model that does not account for land use differences and present the results in Tables A2 and A3 of the online appendix. The results present an inconsistent picture of overall approval: Total WTP switches signs across specifications. These results suggest that failure to adequately control for current land use can obscure strong preferences for and against specific types of solar development, thus underscoring the importance of including current land use in the discussion surrounding the siting of solar installations.

5 POLICY RECOMMENDATIONS

In order to stimulate solar growth and achieve renewable energy targets, Rhode Island buys renewable energy from producers at a premium to offset the higher levelized cost than conventional sources. However, the incentives offered to solar developers are constant regardless of the attributes of the project. Given the additional costs of developing on commercial/industrial areas, brownfields, and covered landfills, the constant incentive essentially encourages solar development on farm and forest lands. In addition, visual barriers from landscaping or other means are additional costs to developers, and thus may be insufficiently provided.

Several New England and Mid-Atlantic states do offer differentiated subsidies based on prior land use (see Knight et al. 2020 for a review). The most common is an additional incentive for parking lot canopies. For example, Massachusetts offers an additional $0.06/kWh and Maryland offers up to $400 per kW of installed capacity. Rhode Island undertook a pilot project in 2020 offering a $0.06/kWh adder for a single solar parking lot canopy development (*RIPUC*, n.d.). Several states similarly offer differentiated rates for solar built on brownfields and covered landfills. In the case of Massachusetts, this is a \$0.03/kWh and \$0.04/kWh adder, respectively. Vermont offers financial resources for assessment and cleanup of contaminated sites. Massachusetts additionally uses disincentives for solar sited on forest land. The deduction increases with the size of the installation, but as an example a 5 MW array would receive a deduction of \$0.015/kWh from the standard incentive (*MA-Smart Solar*, n.d.). While these differentiated incentives are certainly in line with our estimates of preferences across land types, it is unclear whether they pass a benefit-cost test or if similar differentiated incentives enacted in Rhode Island would pass a benefit-cost test.

While the results presented in Section 5 indicate welfare impacts to households from various solar siting decisions, we additionally seek to use our results to inform policy. As illustrated above, many policy actions take the form of per kWh incentives or disincentives, so that is how we structure our analysis here. Table 7 presents the logical steps of converting our household valuation results into per kWh incentives for various policy actions that are costly to developers but preferred by residents (i.e., moving development from forest land to commercial land). The goal is to develop incentives that are justified based on residents' preferences. We conduct this exercise based on a 2 MW array. Column 1 is monthly household WTP for each policy action and is calculated from Column 3, Panel B of Table 5. Column 2 is this household WTP per kWh of production. This equals Column 1 divided by 237,600 kWh, which is expected monthly electricity generation from a 2 MW array with a capacity factor of 16.5%.

The remaining columns aggregate WTP across households within a given distance (0.5, 1, 3 miles) of a hypothetical solar array. The number of households within a given distance is approximated using census data for the whole state of Rhode Island. We present multiple distances because it is uncertain what the appropriate aggregation level is. A distance of 0.5 miles might approximate the size of an area in which residents are likely to frequently encounter a solar array. Another measure of proximity stems from two studies that find that property value impacts extend to about one mile: Gaur and Lang (2020) in Massachusetts and Rhode Island, and Abashidze (2019) in North Carolina. Often solar developments are hotly debated at town meetings, and the average town in Rhode Island has an approximate radius of three miles, so we present that as an upper bound.

The results suggest that, even for conservative definitions of impacted households, substantial incentives are justified. For example, aggregating over only residents within 0.5 miles, an additional incentive of $0.07/kWh is justified if a solar array development is moved from forest land to commercial land. Similarly, an additional incentive of $0.06/kWh is justified if a solar array development is moved from forest land to a brownfield. Incentives to displace development on farmland are smaller at $0.03/kWh. Incentives for visibility screening come in around $0.01/kWh. As the distance of impacted households grows, so do the incentives justified, reaching excessive levels for this context (i.e., $2.47/kWh for moving a development from forest to commercial).

These incentives can additionally be altered to reflect the reality of development proposals. For instance, a developer cannot credibly declare they would build on forest land, but are now building on a brownfield, and so deserve a

$0.06/kWh added incentive. One option would be to place a $0.03/kWh added incentive on brownfields, and a reduced feed-in-tariff of $0.03/kWh if an array is sited on forest lands. This combination would mirror resident preferences for land types. When it comes to screening, landscaping typically is an upfront fixed cost, and thus would not need an ongoing per kWh incentive. However, vegetative (or even artificial) buffers can deteriorate over the 25 year lifetime of an array if not tended, thus an annual verification of visual screening to qualify for a small incentive (per kWh or a flat fee) could be appropriate.

As mentioned above, our calculations in Table 7 use a 2 MW capacity. As capacity grows, production grows, and subsidies decline. Since household WTP values are independent of any assumptions of solar attributes, only electricity generation will be affected when we assume an installation with a different capacity. Therefore, the WTP/kWh values will decrease in proportion to the size of the assumed installation. In Appendix Table A4, we present an analogous version of Table 6 using a 6 MW capacity installation instead. Justified incentives are substantially smaller, however, this may be appropriate as levelized cost goes down as capacity increases (*RIPUC*, n.d.).

6 CONCLUSION

This paper quantifies the externalities of utility-scale solar installations by analyzing RI residents' tradeoffs for six solar siting attributes: size of the installation, visibility, setback distance, probability of future residential development, change in electricity bill, and current land use of the proposed solar site. We collect data using a survey that was distributed to a random sample of 2,794 RI residents. Our final sample consists of 3,936 choices made by 656 respondents.

We use a CE framework and logistic regression models to estimate respondents' WTP for each attribute. MWTP values indicate that Rhode Islanders like large installations and are willing to pay \$0.28 for each additional acre of land to be developed for solar energy. However, respondents dislike fully visible installations and need to be compensated \$8.43 for the same. We find no significant impacts from setback distance and partly visible installations, suggesting that respondents are unaffected by these attributes. When the probability of future residential development increases, they are less likely to choose the status quo alternative of no solar development.

Assuming a 10 acre, fully visible installation with a setback distance of 150 feet and 0% probability of future residential development, we obtain total WTP values for solar development on different land types. Our results indicate substantial heterogeneity in preferences for constructing solar installations by current land use of a proposed solar site. Overall, respondents dislike solar development on farmlands and forests, and need to be compensated \$13 to \$49 per month for the change. However, they support solar development on brownfields and commercial land types and are willing to pay an additional \$15 to \$19 per month to have solar installations constructed there. It is important to remember that our sample respondents overwhelmingly have positive attitudes towards solar energy. Our results provide nuance to that favorability. Concerns heard about solar developments in town meetings and stakeholder groups are not likely blanket NIMBY concerns, but instead are concerns about land use change and other important priorities.

We conclude with calculations and a discussion about how our results can be converted to policy relevant parameters. The incentives and disincentives will promote solar development that is consistent with residents' preferences. As Rhode Island and other states seek to meet renewable energy objectives, assessment and incorporation of residents' preferences are critical to ensure ongoing support.

This research extends the literature on both preferences for utility-scale renewable electricity generation and preferences for land conservation in a manner that is relevant to stakeholders and residents as well as actionable for policy makers. We model land use in the solar siting choice in a way that is intuitive, clear, and obviously resonant to Rhode Islanders. We also extend our analysis beyond traditional household WTP estimates to frame resident preferences in a way that mirrors the units of subsidy for utility scale installations. This research also hints at important future extensions in this area. Spatial heterogeneity of preferences likely exists in this area and has been shown relevant to preferences for land conservation in other contexts (Czajkowski et al., 2016). Modeling how preferences vary by spatial distance is an important extension of this work and will help shed light on which spatial aggregation ranges from Table 7 are most appropriate when determining incentives.

Figures and tables

Figure 1: Example choice question

Parcel 1: Farmland

Consider a group of privately-owned land parcels that totals 50 acres and are currently used as farmland. These parcels are in Rhode Island and less than 15 minutes from your residence by car. Below are two possible solar development plans for these farmland parcels. Policy makers can approve either plan, or they can reject both plans and have no solar installation on the parcels.

Please examine the three options below and indicate which option you prefer.

	CHOICE A	CHOICE B	CHOICE C
Size of installation	10 acres (generates enough power for 320 homes)	30 acres (generates enough power for 960 homes)	NO SOLAR PANELS
Visibility	Visible	Not visible	
Setback	100 ft	50 ft	
Probability of residential development	0%	0%	50%
Change in monthly electricity bill (annual)	$10 increase ($120 annual ↑)	$10 decrease ($120 annual ↓)	No change
YOUR CHOICE [Please check ONE box only]	**A** ☐	**B** ☐	**C** ☐

Table 1: Attribute definitions and levels

Attribute	Definition	Levels
Size of installation	The size of the solar installation in acres.	10, 20, 30, 40, 50
Visibility	Visibility of a solar installation from a respondent's house or from regularly traveled roads.	Not visible, Partially visible, Completely visible
Setback[a]	Minimum distance of the solar panels from the property line.	0, 50, 100, 250
Probability of residential development[b]	The likelihood that the land being considered will be developed into residential housing in the next ten years if a solar installation is not built.	0%, 25%, 50%
Change in electricity bill[c]	The dollar increase or decrease in a respondent's monthly electricity bill if the parcel is converted to solar power generation.	-$30, -$20, -$10, -$5, $5, $10, $20, $30
Current land use[d]		
a) Farmland	The land is currently used to grow agricultural crops. In this case, solar installations would be built on the ground.	
b) Forest	The land is currently privately-owned forest land. In this case, trees will be clear cut and solar installations would be built on the ground.	
c) Commercial	The land is either currently used for business activities, including buildings and parking lots, or undeveloped land that is zoned for commercial purposes. In this case, solar installations could be built on the ground, on building rooftops, or as a parking lot canopy.	
d) Brownfield	A former industrial or commercial site where future use is affected by real or perceived environmental contamination. These could include capped landfills and quarries. In this case, solar installations would be built on the ground.	

Notes: [a] Setback level of 0 feet is excluded for farm and forest land use types.

[b] Probability of residential development is excluded when the land use type is commercial or brownfield.

[c] For the commercial and brownfield land types, the levels of -$30, -$20, and -$10 are excluded.

[d] Current land varies across choice tasks but is constant across alternatives within each choice task.

Table 2: Summary statistics of survey respondents

Variable	Mean	SD	Minimum	Maximum	Observations
Household income (000's)	109.25	50.96	15	175	601
College educated (1 = yes)	0.68	0.47	0	1	649
Children at home (1 = yes)	0.35	0.48	0	1	646
Female (1 = yes)	0.52	0.50	0	1	656
Homeowner (1 = yes)	0.83	0.38	0	1	647
Years living in current home	15.51	6.55	3	20	651
Employed (1 = yes)	0.63	0.48	0	1	650
Electricity bill ($/month)	123.57	54.88	25	200	646
Solar panels at home (1 = yes)	0.05	0.23	0	1	647
Energy attitudes (1 = positive)					
Solar	0.87	0.33	0	1	649
Offshore wind	0.80	0.40	0	1	640
Onshore wind	0.76	0.43	0	1	638
Hydro	0.68	0.47	0	1	637
Natural gas	0.52	0.50	0	1	642
Nuclear	0.26	0.44	0	1	627
Coal	0.08	0.28	0	1	638
Consequentiality (1 = positive)					
Policy	0.91	0.29	0	1	656
Payment	0.90	0.29	0	1	656

Notes: All data come from survey responses. Household income and electricity bill values come from a multiple choice question that included several ranges. We assign people the middle value of their chosen range.

Table 3: Summary statistics for sampling weighting variables

		Unweighted sample	Population	Weighted sample
Location	Rural (%)	73.32	50.62	50.62
	Urban (%)	26.68	49.38	49.38
Age	18 - 39 (%)	23.93	33.69	33.69
	40 - 59 (%)	33.84	31.83	31.83
	60+ (%)	42.23	34.49	34.49
Party	Democrat (%)	34.76	39.73	39.73
	Republican (%)	19.21	12.20	12.20
	Independent (%)	46.04	48.07	48.07
Number of observations		656	778,666	656

Notes: Data come from Rhode Island voter registration database. All values are represented as percentages of the total number of observations.

Table 4: Attribute coefficients from logit regressions

Variable	Conditional Logit	Heteroscedastic Logit	Random Parameters logit	
			Mean	SD
Acres	0.010***	0.018***	0.016***	0.044***
	(0.002)	(0.004)	(0.003)	(0.004)
PartVisibility	-0.066	-0.067	-0.127	0.154
	(0.061)	(0.085)	(0.083)	(0.279)
FullVisibility	-0.313***	-0.406***	-0.546***	0.801***
	(0.074)	(0.103)	(0.111)	(0.196)
Setback (00's ft)	0.042	0.066	0.079*	0.293***
	(0.031)	(0.042)	(0.044)	(0.109)
Probability	-0.008***	-0.014***	-0.029***	0.077***
	(0.002)	(0.004)	(0.007)	(0.014)
Cost ($/month)	-0.043***	-0.065***	-0.065***	
	(0.002)	(0.007)	(0.004)	
Land use ASC interactions				
Farm × ASC	0.822***	1.407***	0.590*	3.859***
	(0.135)	(0.268)	(0.333)	(0.600)
Forest × ASC	1.596***	2.988***	2.910***	4.161***
	(0.134)	(0.465)	(0.374)	(0.560)
Brownfield × ASC	-0.793***	-0.782***	-1.232***	0.338
	(0.128)	(0.148)	(0.165)	(0.217)
Commercial × ASC	-1.035***	-1.068***	-1.517***	0.045
	(0.132)	(0.190)	(0.167)	(0.220)
Heteroskedastic variables				
Farm		-0.450***		
		(0.122)		
Forest		-0.704***		
		(0.142)		
Commercial		-0.026		
		(0.138)		
Choices	11,724	11,724	11,724	
Respondents	656	656	656	
AIC	7389.010	7347.585	6615.204	
BIC	7462.704	7443.387	6755.223	

Note: Acres refers to the size of the solar installation in acres. Part visibility and Full visibility are dummy variables = 1 if a solar installation is partially or completely visible, respectively. ASC is the status-quo alternative-specific constant, or a dummy variable = 1 for the status-quo choice and 0 otherwise. Cost is in terms of USD per household per month. Sample weights are applied and constructed using stepwise adjustment on three variables: age, political affiliation, and rural/urban residence. Standard errors, clustered by respondent, are in parentheses. *,**,*** indicate significance at the 90%, 95% and 99% level, respectively.

Table 5: Marginal willingness to pay estimates for solar attributes

Attribute	Conditional logit	Heteroskedastic logit	Random parameters logit
Panel A: Marginal WTP			
Acres	$0.24***	$0.28***	$0.25***
	(0.05)	(0.05)	(0.05)
PartVisibility	-$1.54	-$1.03	-$1.96
	(1.44)	(1.34)	(1.27)
FullVisibility	-$7.30***	-$6.21***	-$8.43***
	(1.75)	(1.50)	(1.63)
Setback (00's ft)	$0.98	$1.01	$1.21*
	(0.71)	(0.62)	(0.68)
Probability	-$0.19***	-$0.22***	-$0.45***
	(0.06)	(0.07)	(0.11)
Panel B: Total WTP			
Farmland	-$22.54***	-$23.43***	-$13.22***
	(2.67)	(2.63)	(5.01)
Forest	-$40.58***	-$47.62***	-$49.04***
	(2.76)	(4.08)	(5.36)
Commercial	$20.72***	$14.43***	$19.32***
	(3.09)	(2.45)	(2.71)
Brownfield	$15.07***	$10.06***	$14.91***
	(2.89)	(2.29)	(2.58)

Notes: Welfare estimates are in USD per household per month. Estimates in Panel A represent marginal WTP values. In Panel B, the estimates represent total WTP values and assume a 10 acre, fully visible installation with a setback of 150 feet, and a 0% probability of development in the future. In both panels, standard errors are calculated using the bootstrap method (with 1000 replications) and are in parentheses. *,**,*** indicate significance at the 90%, 95% and 99% level, respectively.

Table 6: Marginal willingness to pay estimates for solar attributes estimated separately by land types

Attribute	Farm and Forest (1)	Commercial and Brownfield (2)
Panel A: Marginal WTP		
Acres	$0.08	$0.38***
	(0.07)	(0.06)
PartVisibility	-$3.38*	$0.28
	(1.99)	(1.81)
FullVisibility	-$10.34***	-$4.42**
	(2.68)	(1.74)
Setback (00's ft)	$0.18	$1.35*
	(1.09)	(0.84)
Probability	-$0.21***	
	(0.06)	
Panel B: Total WTP		
Farmland	-$21.63***	
	(3.64)	
Forest	-$41.63***	
	(3.91)	
Commercial		$13.49***
		(2.41)
Brownfield		$9.30***
		(2.25)

Notes: Welfare estimates are in USD per household per month. Estimates in Panel A represent marginal WTP values. In Panel B, the estimates represent total WTP values and assume a 10 acre, fully visible installation with a setback of 150 feet, and a 0% probability of development in the future. In both panels, standard errors are calculated using the bootstrap method (with 1000 replications) and are in parentheses. *,**,*** indicate significance at the 90%, 95% and 99% level, respectively.

Table 7: Developing solar siting incentives justified by residents' preferences

Policy Action	Household WTP	Household WTP/kWh	Aggregate WTP/kWh		
			Median households within 0.5 miles	Median households within 1 mile	Median households within 3 miles
	(1)	(2)	(3)	(4)	(5)
Forest to Commercial	\$68.36	\$0.00029	\$0.07	\$0.27	\$2.47
Forest to Brownfield	\$63.95	\$0.00027	\$0.06	\$0.26	\$2.31
Farm to Commercial	\$32.54	\$0.00014	\$0.03	\$0.13	\$1.18
Farm to Brownfield	\$28.13	\$0.00012	\$0.03	\$0.11	\$1.02
Fully visible to partly visible	\$6.47	\$0.00003	\$0.01	\$0.03	\$0.23
Fully visible to not visible	\$8.43	\$0.00004	\$0.01	\$0.03	\$0.30

Notes: Household WTP values in Column 1 are derived from Column 3 of Table 5. The WTP for switching solar development from one land type to another is calculated by subtracting the total WTP for the former land type from the latter. The WTP for converting a fully visible installation into a partly visible one is obtained by subtracting the WTP for a partly visible installation from the WTP for a fully visible installation, and then changing the sign from negative to positive. The WTP for making a fully visible installation not visible at all is the negative of the marginal WTP estimate of $FullVisibility$. Column 2 values are calculated by dividing Column 1 by expected monthly electricity generation from a 2 MW installation. Columns 3, 4, and 5 take the household WTP/kwh values from Column 2 and aggregate them over the median number of households within a radius of 0.5, 1, 3 miles respectively. Based on population density from the 2010 RI Census, we calculate the median number of households within an area equivalent to 0.5 miles from a solar array is 239, within 1 mile is 955, and within 3 miles is 8,599.

Appendix

This appendix provides supplemental figures and tables to our main results.

Figures A1, A2, and A3 depict example choice questions for the forest, commercial, and brownfield land types, respectively.

Figure A4 presents the proportion of respondents that chose the status quo and the solar development plans across different types of land use. For forested land, more respondents chose the status quo option over solar development. For brownfield and commercial land types, a large majority (almost 80%) choose solar development over the status quo option.

Figure A5 represents the importance of each attribute for the respondents while making their decisions. Land use is the most important, which underscores the importance of including current land use in the context of solar siting. $Setback$ is the least important, which explains the insignificant coefficient associated with it across all models.

Table A1 presents conditional logit regression estimates derived by estimating Equation (8) on two types of lands: desirable (farmland and forest), and undesirable (commercial and brownfield) in Columns 1 and 2, respectively. We find that the coefficient on $Acres$ is insignificant for farm and forest land types, but highly significant (at the 1% level) and positive for commercial and brownfield land types, implying that respondents prefer larger solar installations only on undesirable land types. The coefficient on $PartVisibility$ for the farm and forest column is negative, suggesting that responds dislike partly visible installations on those lands, though the coefficient is significant only at the 10% level. However, they have no such preference for installations on commercial and brownfield lands (positive and insignificant coefficient). Fully visible installations are disliked regardless of current land use, implying that solar installations cause significant visual disamenities wherever they may be sited. The coefficient on $Setback$ is positive and significant (at the 10% level) for commercial and brownfield land types only. This could represent respondents' willingness to remain distant from undesirable land types. The coefficient on $Probability$ is negative and significant at the 1% level, similar to our main results, implying that as the probability of future residential development increases, people become less likely to choose the status quo option. The positive and significant coefficients on the $Farm \times ASC$ and $Forest \times ASC$ interaction terms, and the negative and significant ones on the $Brownfield \times ASC$ and $Commercial \times ASC$ terms are in line with our main results: that respondents dislike solar installations on farms and forest lands but support their construction on commercial and brownfield lands.

Table A2 presents the coefficients derived from estimating a basic model with no land use-ASC interaction terms. The coefficients on $Acres$, $PartVisibility$, and $FullVisibility$, are qualitatively similar to our main results, with respondents demonstrating a liking for large installations that are not visible. The sign of the $Setback$ coefficient is inconsistent across models, though it is insignificant throughout implying that the average respondent is unaffected by setback distance. $Probability$ is significant and positive, which is in contrast to our main results, but makes sense in this context. Since this model does not differentiate between land use, the only way respondents' preference for keeping the status-quo for farm and forest lands can be captured is by having a positive coefficient for $Probability$, which takes positive values only when the land use is forest or farmland. The ASC coefficient is inconsistent across models in both sign and significance. It is positive and insignificant in the CL model, negative and significant in the HCL and RPL models. The negative sign indicates that respondents prefer to choose a solar development plan over maintaining the status quo. This result is not surprising, given that over 80% of the subjects in our sample support solar energy. However, the large and significant SD value in the RPL model indicates that there is some heterogeneity in their preferences that remains unexplained, which we examine in our main specification with land use interactions.

Table A3 presents welfare estimates derived from our basic model with delta standard errors. As with our main model, we consider a 10 acre solar installation with full visibility, 150 feet setback distance and 0% probability of future residential development. The MWTP values for $Acres$, $PartVis$ibility, and $FullVisibility$, are quite similar in both sign and magnitude compared to our main results. For $Setback$, the MWTP is negative for the CL and the RPL models, and positive for the HCL model, though insignificant throughout. We find that respondents' MWTP for $Probability$ is positive and significant, a result that is opposite to our main results. Finally, without accounting for the differences in land use, we find that our CV estimates are inconsistent across different models. The CL estimate is negative and significant at the 1% level, indicating that respondents need to be compensated \$8.29 for the construction of the particular solar installation under consideration. However, the CLH estimate implies a positive WTP of \$4.27 per month for the same kind of installation (though it is insignificant). The RPL value is negative and insignificant. Not only do these results underestimate the compensation levels for when construction happens on forest lands, but also the respondents' positive WTP for solar development on commercial land and brownfields.

Table A4 replicates table 6 of our main manuscript but assumes electricity generation from a 6 MW installation. The incentives are smaller than our main results for all levels of aggregation but may still be realistic since levelized costs go down as capacity increases.

Figure A1: Example choice question for forest land parcels

Parcel 2: Forest

Consider a group of privately-owned forested land parcels that total 50 acres and are currently undeveloped. These parcels are in Rhode Island and less than 15 minutes from your residence by car. Below are two possible solar development plans for these forest parcels. Policy makers can approve either plan, or they can reject both plans and have no solar installation on the parcels.
Please examine the three options below and indicate which option you prefer.

	CHOICE A	CHOICE B	CHOICE C
Size of installation	20 acres (generates enough power for 640 homes)	50 acres (generates enough power for 1,600 homes)	NO SOLAR PANELS
Visibility	Completely visible	Partially visible	
Setback	100 ft	50 ft	
Probability of residential development	0%	0%	50%
Change in monthly electricity bill (annual)	$5 increase ($60 annual ↑)	$15 decrease ($180 annual ↓)	No change
YOUR CHOICE [Please check ONE box only]	**A** ☐	**B** ☐	**C** ☐

Figure A2: Example choice question for commercial land parcels

Parcel 3: Commercial land

Consider a group of privately-owned land parcels that total 50 acres and are currently used or zoned as commercial land. These parcels are in Rhode Island and less than 15 minutes from your residence by car. Below are two possible solar development plans for these commercial parcels. Policy makers can approve either plan, or they can reject both plans and have no solar installation on the parcels.

Please examine the three options below and indicate which option you prefer.

	CHOICE A	**CHOICE B**	**CHOICE C**
Size of installation	30 acres (generates enough power for 960 homes)	50 acres (generates enough power for 1,600 homes)	NO SOLAR PANELS
Visibility	Partially visible	Completely visible	
Setback	50 ft	100 ft	
Change in monthly electricity bill (annual)	$15 increase ($180 annual ↑)	$20 increase ($240 annual ↑)	No change
YOUR CHOICE [Please check ONE box only]	**A** ☐	**B** ☐	**C** ☐

Figure A3: Example choice question for brownfield land parcels

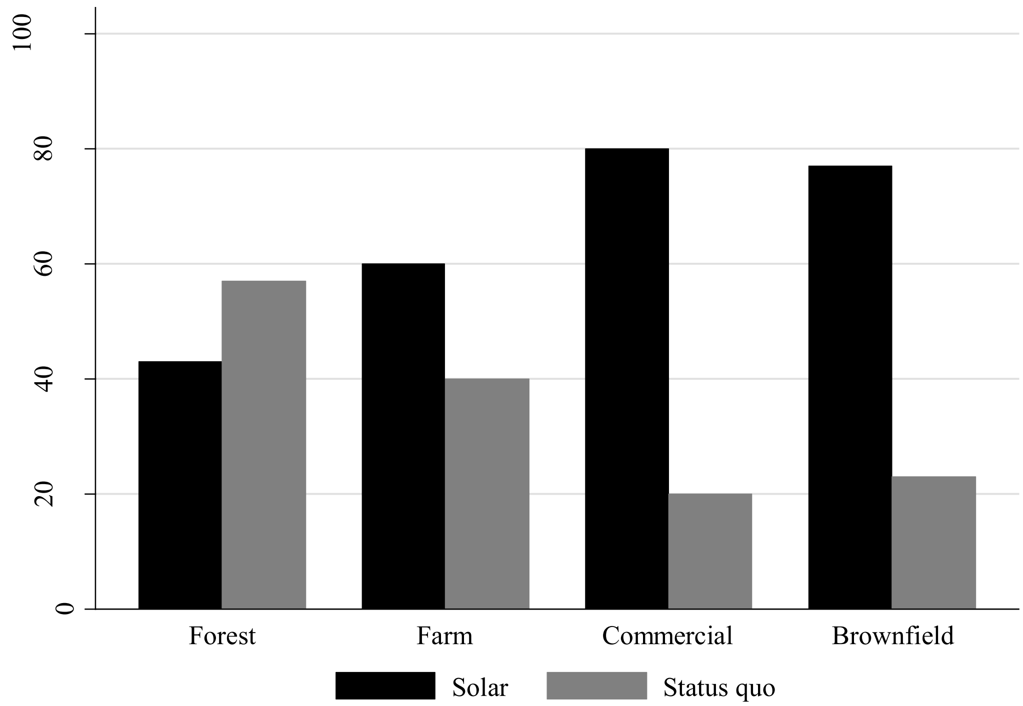
Parcel 4: Brownfield

Consider a group of privately-owned land parcels that total 50 acres and are currently brownfields. These parcels are in Rhode Island and less than 15 minutes from your residence by car. Below are two possible solar development plans for these brownfield parcels. Policy makers can approve either plan, or they can reject both plans and have no solar installation on the parcels.

Please examine the three options below and indicate which option you prefer.

	CHOICE A	CHOICE B	CHOICE C
Size of installation	20 acres (generates enough power for 640 homes)	30 acres (generates enough power for 960 homes)	NO SOLAR PANELS
Visibility	Partially visible	Not Visible	
Setback	100 ft	200 ft	
Change in monthly electricity bill (annual)	$10 increase ($120 annual ↑)	$15 increase ($180 annual ↑)	No change
YOUR CHOICE [Please check ONE box only]	**A** ☐	**B** ☐	**C** ☐

Figure A4: Alternative choice by land use



Notes: N = 1,298 for forest, 1,305 for Farm, 652 for commercial, and 653 for brownfield.

Figure A5: Importance of attributes while making choices

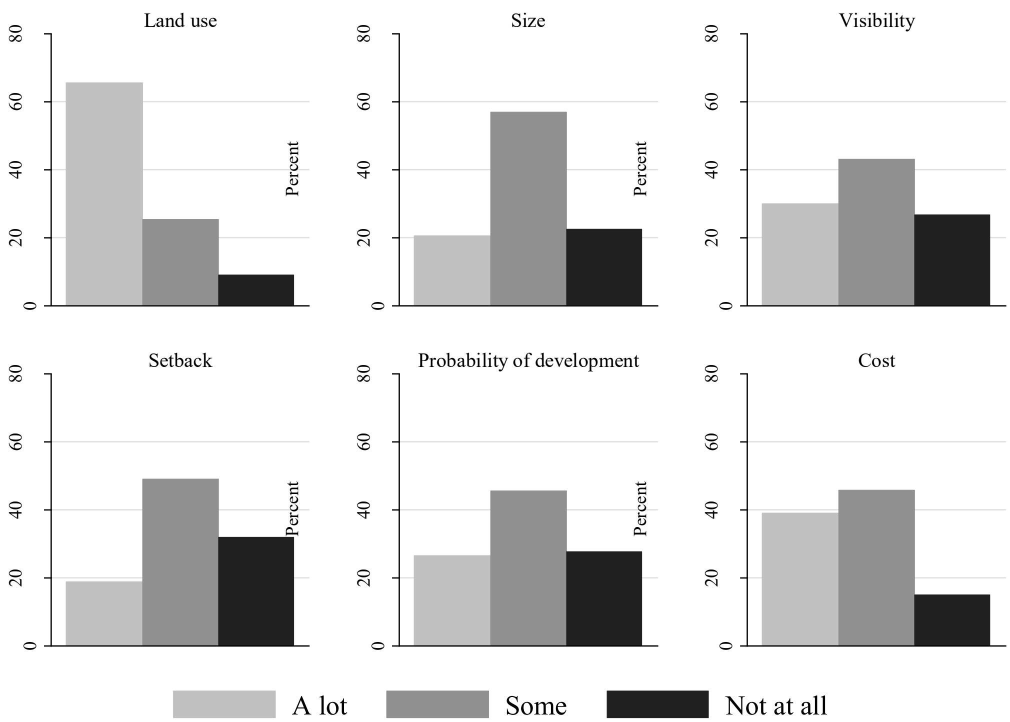


Table A1: Attribute coefficients estimated separately by land types

Variable	Farm and Forest (1)	Commercial and Brownfield (2)
Acres	0.003	0.024***
	(0.003)	(0.004)
PartVisibility	-0.130*	0.018
	(0.077)	(0.111)
FullVisibility	-0.397***	-0.280**
	(0.098)	(0.113)
Setback (00's ft)	0.007	0.086*
	(0.041)	(0.050)
Probability	-0.008***	
	(0.002)	
Cost ($/month)	-0.038***	-0.063***
	(0.002)	(0.005)
Interactions		
Farm × ASC	0.476***	
	(0.138)	
Forest × ASC	1.244***	
	(0.136)	
Brownfield × ASC		-0.501***
		(0.167)
Commercial × ASC		-0.767***
		(0.169)
Observations	7,809	3,915
AIC	4915.889	2421.016
BIC	4971.593	2464.924

Note: All estimates are derived from conditional logit regressions. Acres refers to the size of the solar installation in acres. PartVisibility and FullVisibility are dummy variables = 1 if a solar installation is partially or completely visible, respectively. ASC is the status-quo alternative-specific constant, or a dummy variable = 1 for the status-quo choice and 0 otherwise. Cost is in terms of USD per household per month. Sample weights are applied and constructed using stepwise adjustment on three variables: age, political affiliation, and rural/urban residence. Standard errors, clustered by respondent, are in parentheses. *,**,*** indicate significance at the 90%, 95% and 99% level, respectively.

Table A2: Attribute coefficients from logit regressions for basic specification without prior-land-use-specific ASC

Variable	Conditional Logit	Heteroscedastic Logit	Random parameters logit	
			Mean	SD
Acres	0.010***	0.024***	0.012***	0.020***
	(0.002)	(0.004)	(0.002)	(0.004)
PartVisibility	-0.046	-0.029	-0.081	0.004
	(0.060)	(0.105)	(0.067)	(0.068)
FullVisibility	-0.276***	-0.418***	-0.372***	0.428**
	(0.070)	(0.117)	(0.085)	(0.200)
Setback (00's ft)	-0.026	0.032	-0.031	0.184*
	(0.030)	(0.050)	(0.039)	(0.105)
Probability	0.016***	0.061***	0.024***	0.046***
	(0.002)	(0.013)	(0.003)	(0.008)
ASC	0.053	-0.434**	-0.236**	1.274***
	(0.096)	(0.153)	(0.119)	(0.121)
Cost ($/month)	-0.033***	-0.071***	-0.039***	
	(0.002)	(0.007)	(0.003)	
Heteroskedastic variables				
Farm		-1.063***		
		(0.206)		
Forest		-1.263***		
		(0.153)		
Commercial		-0.005		
		(0.130)		
Choices	11,724	11,724	11,724	
Respondents	656	656	656	
AIC	7967.51	7880.30	7468.17	
BIC	8019.10	7954.00	7563.97	

Note: Acres refers to the size of the solar installation in acres. Part visibility and Full visibility are dummy variables = 1 if a solar installation is partially or completely visible, respectively. ASC is the status-quo alternative-specific constant, or a dummy variable = 1 for the status-quo choice and 0 otherwise. Cost is in terms of USD per person per month. Standard errors, clustered by respondent, are in parentheses. *,**,*** indicate significance at the 90%, 95% and 99% level, respectively.

Table A3: Welfare estimates for solar attributes without prior-land-use-specific ASC

Attribute	Conditional logit	Heteroskedastic logit	Random parameters logit
Acres	$0.30***	$0.34***	$0.29***
	(0.06)	(0.07)	(0.06)
PartVisibility	-$1.42	-$0.41	-$2.05
	(1.85)	(1.59)	(1.70)
FullVisibility	-$8.47***	-$5.92***	-$9.46***
	(2.21)	(1.79)	(2.12)
Setback (00's ft)	-$0.80	$0.45	-$0.78
	(0.89)	(0.67)	(0.91)
Probability	$0.50***	$0.87***	$0.60***
	(0.06)	(0.24)	(0.09)
Total WTP	-$8.29***	$4.27	-$1.43
	(2.75)	(6.45)	(2.81)

Notes: Welfare estimates are in USD per person per month. The total WTP estimates assume a 10 acre, fully visible installation with a setback of 150 feet, and a 0% probability of development in the future. Standard errors, calculated using the bootstrap method (1000 replications) , are in parentheses. *,**,*** indicate significance at the 90%, 95% and 99% level, respectively.

Table A4: Developing solar siting incentives justified by residents' preferences, 6 MW array

Policy Action	Household WTP	Household WTP/kWh	Aggregate WTP/kWh: Median households within 0.5 miles	Aggregate WTP/kWh: Median households within 1 mile	Aggregate WTP/kWh: Median households within 3 miles
	(1)	(2)	(3)	(4)	(5)
Forest to Commercial	$68.36	$0.00029	$0.03	$0.09	$0.82
Forest to Brownfield	$63.95	$0.00027	$0.03	$0.09	$0.77
Farm to Commercial	$32.54	$0.00014	$0.01	$0.04	$0.39
Farm to Brownfield	$28.13	$0.00012	$0.01	$0.04	$0.34
Fully visible to partly visible	$6.47	$0.00003	$0.002	$0.01	$0.08
Fully visible to not visible	$8.43	$0.00004	$0.003	$0.01	$0.10

Notes: Household WTP values in Column 1 are derived from Column 3 of Table 5. The WTP for switching solar development from one land type to another is calculated by subtracting the total WTP for the former land type from the latter. The WTP for converting a fully visible installation into a partly visible one is obtained by subtracting the WTP for a partly visible installation from the WTP for a fully visible installation, and then changing the sign from negative to positive. The WTP for making a fully visible installation not visible at all is the negative of the marginal WTP estimate of $FullVisibility$. Column 2 values are calculated by dividing Column 1 by expected monthly electricity generation from a 6 MW installation. Columns 3, 4, and 5 take the household WTP/kwh values from Column 2 and aggregate them over the median number of households within a radius of 0.5, 1, 3 miles respectively. Based on population density from the 2010 RI Census, we calculate the median number of households within an area equivalent to 0.5 miles from a solar array is 239, within 1 mile is 955, and within 3 miles is 8,599.

Manuscript – 3

Prepared for submission to Applied Economics Letters

Do the Loudest Voices in the Room Speak for All of Us? Comparing Engaged Versus Random Samples in a Contingent Valuation Framework

Vasundhara Gaur [a], Corey Lang [a], Gregory Howard [b], and Ruth Quainoo [b]

[a] Department of Environmental and Natural Resource Economics, University of Rhode Island
[b] Department of Economics, East Carolina University

ABSTRACT
This study estimates and compares willingness to pay values from a choice experiment regarding utility-scale solar energy between two groups: a traditional, random sample derived using best practices and a sample of engaged individuals who attended an informational meeting about solar energy. Our results suggest that the preferences of engaged and random sample respondents align in sign for a majority of attributes. However, there are large differences in willingness to pay magnitudes, with engaged respondents willing to pay two to four times more than random sample respondents. We advise caution while generalizing valuation estimates derived from convenience samples, though the overall preferences can still provide insights of the broader population.

Keywords: Convenience sample; Contingent valuation; Utility-scale solar; Willingness to pay
JEL codes: C83; Q24; Q42; Q51

1 INTRODUCTION

This paper is motivated by one policy question and a related academic question. Important local policy decisions about school funding, land development, and budget priorities are often decided by city or town councils, with input from school committees, zoning boards, and similar. In addition, these councils and committees hold public meetings, in which residents can attend and speak their opinions on issues under consideration. In *Democracy in America*, de Tocqueville viewed township democracy as a particularly strong pillar of the American experiment. However, at least in its modern incarnation, only a select few residents show up and speak at town meetings, and often they are the most zealous for a given issue. This raises the concern that local decisions are made with input from people whose preferences are not representative. This paper investigates how preferences for an engaged group of residents differ from a representative sample, and in turn how following policy prescriptions from those who show up to meetings may lead to policy that does not reflect the whole town.

The academic motivation for this paper is assessing how convenience sampling may yield biased results when seeking to estimate population preferences. Random sampling is the gold standard of collecting data, but it can often be costly and cumbersome to collect such a sample. In contrast, convenience sampling is cheaper and easier, but it is unclear to what extent inferences drawn from that kind of sample are generalizable. To this end, we investigate differences in preferences and willingness to pay (WTP) values between a random sample and a convenience sample comprised of engaged stakeholders.

We design a choice experiment (CE) survey to elicit WTP values for solar siting attributes in the state of Rhode Island. Four solar siting attributes are considered: installation size, visibility, setback distance, and current land use of the proposed development site. To recruit highly engaged stakeholders, a webinar titled "Valuing Siting Options for Solar Energy in RI" was organized by the University of Rhode Island and advertised on social media. The content of the webinar discussed the trends in solar development and presented an overview of research plans, but no findings or opinions were given. Following the webinar, the survey was disseminated to the list of

104 people who had registered for the seminar. Of those, 45 people returned complete responses and form our engaged sample. In addition, a random sample of 3000 individuals were drawn from the 2019 Rhode Island voter registration database and invited to take the survey. Of those, 656 completed our survey and form our random sample of respondents.

Our results indicate that the engaged and the random sample respondents' preferences lie in the same direction for the most part (in that they like and dislike the same set of attributes) but differ in magnitude. As a result, large differences exist in some of their respective WTP values. Both groups of respondents like large installations, and the engaged sample willing to pay almost twice as much (between \$12 and \$17.40 per month) for the same. To avoid visual disamenities from solar installations, respondents in the random sample are willing to pay between \$6.21 and \$8.43 per month, and engaged respondents are willing to pay three times more than that, between \$22.12 and \$27.28 per month. While random sample respondents are mainly unaffected by setback distance, the engaged sample care much more about it. Our results also indicate that when the probability of future residential development on farmland and forest increases, random sample respondents are less likely to choose the status quo option. In contrast, engaged respondents are not affected at all by future residential development on farms and forests. The random sample respondents are also less sensitive to changes in their utility bill.

We find that both groups dislike having farms and forests converted into solar installations. Engaged respondents are willing to pay between \$82.08 and \$198.50 per month for avoiding solar development on forest lands, which is two to four times more than the random sample's monthly WTP (between \$40.58 and \$49.49). Random sample respondents are in support of developing commercial and brownfield sites into solar, with a total WTP value between \$14.31 and \$20.72 per month for solar development on commercial land, and between \$9.97 and \$15.07 for brownfield development. Engaged respondents are similarly in favor of converting commercial and brownfield lands into solar but are willing to pay over three times more for commercial development, between \$43.71 and \$73.88 per month, and over two times more for brownfield development, between \$18.90 and \$57.73.

We contribute to the environmental economics literature that examines the differences between a random and a conveniently obtained (non-random) sample. Systematic comparisons of results drawn from random and convenience samples are often found in medical literatures (Hedt and Pagano, 2011; Hultsch et al., 2002; Jannink et al., 1995; Özdemir et al., 2011), but its application in environmental economics is less common. We are aware of only one study by Whitehead (1991) that analyzes the differences between a random and a non-random sample in the context of environmental valuation. He uses a dichotomous choice contingent valuation (CV) method to elicit WTP values for wetland protection in Kentucky from a random sample of individuals drawn from telephone directories and a sample of 'interest group' members of the Buckley Hills Audubon Society (BHAS). Comparing WTP values between the two, he finds that the BHAS sample's WTP is over eight times larger than the random sample, which indicates the potential magnitude of self-selection bias. To the best of our knowledge, we are the first to compare WTP values between a random and non-random sample using a CE framework. Similar to Whitehead (1991), our study also finds that the engaged group has higher WTP values for almost all solar attributes, though the magnitude of the difference is smaller.

Our study also contributes to the literature examining self-selection in CV studies. This issue is frequently studied in environmental economics, as individuals with high environmental values are more likely to respond to a survey which biases estimated WTP values upwards (Bowker and Stoll, 1988; Edwards and Anderson, 1987; Loomis, 1987; Mitchell and Carson, 2013; Walsh et al., 1984). Whitehead (1991) finds that on average, the interest group sample is more educated, has a higher income, has more children and has a higher percentage of male respondents compared to the general (random) sample. However, both samples are similar in age. Our findings are mostly similar: the engaged sample is more educated and has more children compared to the random sample, and age is not statistically different between the two samples. Contrary to the findings of Whitehead (1991), we find that on average, our engaged sample has a lower household income and has a higher percentage of female respondents compared to the random sample, though these differences are not statistically significant.

Lastly, our findings yield insights into potential distortions that can be driven into the local policy process by the practice of holding public meetings and the reality that only the loudest voices show up. NIMBY concerns are central to many local land use issues, including renewable energy siting, and our findings suggest that those concerns may be realized or exacerbated when a town council only hears from the engaged population. Jarvis (2021) documents the significant costs of listening to loudest voices in the context of renewable energy siting in the United Kingdom. Specifically, he estimates that the cost of deploying wind projects in the UK increased between 12% – 36% when local officials' refusal to build new renewable energy projects moved development on to less optimal sites.

2 BACKGROUND

The setting of our study is in the state of Rhode Island (RI) where the siting of solar installations has become contentious in recent years (Kuffner, 2018). RI has established a goal of generating 38.5% of statewide energy from renewable sources by 2035, which has brought roughly 400 acres of land in the state under utility-scale solar development (EIA, 2021). Utility-scale solar arrays (sized 1 MW and above) require ten times more land per installed capacity compared to conventional sources, which has raised concerns about land use change in recent years (Denholm and Margolis, 2008; Ong et al., 2013; Trainor et al., 2016). This is especially pronounced in RI, the smallest state in the country, where the construction of utility-scale solar installations puts immense pressure on its already scarce land resources. Additionally, RI is the second most densely populated state in the country, and several installations are built in residential areas with people living in the immediate vicinity of the arrays, which leads to a decline in property prices (Gaur and Lang, 2020).

The solar siting debate in RI is most noticeable at town council and zoning board meetings where solar development plans are proposed, debated, and eventually approved or denied. The largest source of contention is the common practice of siting utility-scale solar on forest and farmlands. While these are the areas where development is cheapest, they offer many amenities to residents. Environmental groups have come down both in support of and opposed to solar expansion due to the difficult tradeoffs of addressing climate change but maintaining ecosystem services and local food production.

The rest of this chapter proceeds as follows. Section 3 outlines the design of our choice experiment and describes the empirical models used in our analysis. Section 4 describes the data and provides summary statistics. Section 5 details the results of our analysis and section 6 provides concluding remarks.

3 METHODS

In this study we use the stated preference (SP) survey methodology for eliciting individuals' preferences and estimating their WTP for various attributes of utility-scale solar installations. Within the SP context, we employ a CE that is developed using a d-efficient design in Stata.

3.1 Choice experiment design

The purpose of the survey is to quantify the perceived externalities from utility-scale solar installations by estimating the tradeoffs people are willing to make for a set of siting attributes. Choice tasks are modelled as solar development plans on four different kinds of land types: farm, forest, commercial, and brownfield (Lang et al., 2021). Each respondent is presented with six choice tasks, two each for farm and forest land types and one each for commercial and brownfield. Within each plan, the respondent is presented with three hypothetical alternatives: A, B, and C. Alternatives A and B describe a solar installation with several attributes that will be developed on the land type under consideration. Alternative C is a status quo option with no solar development that will leave the land in its current state.

All development plans present four solar attributes: *size of installation* (area of land in acres and number of households powered), *visibility* (whether the installation is fully visible, partly visible, or not at all visible), *setback* (minimum distance of the solar panels from the property line), and *change in electricity bill* (in monthly and annual terms). An additional attribute *probability of residential development* is included when the land type is either farmland or forest to account for the future possibility of undeveloped land being converted into residential housing.

Our survey includes four sections. Section 1 provides information on the history of solar development in Rhode Island and the aims and objectives of our study.

Section 2 includes questions about the respondents' use of energy and their attitudes towards various energy sources. In Section 3, we first familiarize the respondents with our CE design, describing all the attributes included in its framework, and then present the six choice sets in a randomized order. The final section gathers information about respondents' perceived consequentiality of the survey, the importance of each attribute in their decision-making, and respondents' demographic characteristics.

3.2 Empirical models

We draw on Random Utility framework (Manski, 1977; McFadden, 1974) where the utility that respondent i derives from alternative m can be broken into an observable and an unobservable (random) component. This can be expressed as:

$$U_{im} = V_{im} + \varepsilon_{im} \tag{1}$$

where U_{im} is the total utility that is a function of observable utility V_{im} and random utility ε_{im}.

We model respondents' choices using three different methods. First, we use McFadden's (1974) multinomial (conditional) logit (CL) model that assumes preference homogeneity and a homoscedastic error structure. The probability of individual i choosing alternative m can be expressed as:

$$P_{im} = \frac{exp(\lambda V_{im})}{\sum_{n=1}^{N} exp(\lambda V_{in})} \tag{2}$$

Where the positive scale factor λ (inversely proportional to the error variance, σ_ε^2) is constant and normalized to unity.

Our second model of choice is the heteroskedastic conditional logit (HCL) (DeShazo and Fermo, 2002; Hensher et al., 1998), which relaxes the CL model's assumption of a homoscedastic error structure by allowing error variances to and scale parameters to vary across individuals and choices. In this model, the scale parameters are represented as a function of $\boldsymbol{Z_{im}}$, a vector of individual- and choice-specific characteristics, and φ, a parameter describing the effect of those characteristics on error variance:

$$\lambda_{im} = exp(\boldsymbol{\varphi Z_{im}}) \tag{3}$$

The probability of individual i choosing alternative m can therefore be expressed as:

$$P_{im} = \frac{exp(\lambda_{im}\boldsymbol{V_{im}})}{\sum_{n=1}^{M} exp(\lambda_{im}\boldsymbol{V_{in}})} \tag{4}$$

Our third and final model is the random parameters (mixed) logit (RPL) model that further relaxes the independence of irrelevant alternatives (IIA) restriction of the CL model and additionally allows for preference heterogeneity. In this model, the utility derived by individual i from alternative m in situation t is:

$$U_{imt} = \boldsymbol{X_{imt}}(\boldsymbol{\beta} + \boldsymbol{\eta_i}) + \varepsilon_{im} = \boldsymbol{X_{imt}}(\boldsymbol{\beta_i}) + \varepsilon_{im} \tag{5}$$

where $\boldsymbol{X_{imt}}$ represents observed attributes, $\boldsymbol{\beta}$ is a vector of coefficients associated with those attributes, and $\boldsymbol{\eta_i}$ is a vector of standard deviation parameters that captures preference heterogeneity. Preference heterogeneity is captured directly in the RPL model through the vector $\boldsymbol{\beta_i}$, which represents how much an individual i deviates from the population mean. The probability of individual $i's$ sequence of choices $[c_1, c_2, \ldots. c_T]$ is described by:

$$P_{i[c_1,c_2,\ldots c_T]} = \int \ldots \int \prod_t^T \left[\frac{exp(\boldsymbol{X_{imt}\beta_i})}{\sum_{n=1}^{M} exp(\boldsymbol{X_{int}\beta_i})}\right] f(\boldsymbol{\beta})d\boldsymbol{\beta} \tag{6}$$

2.3 Estimation

Our main expected utility specification is given as:

$$V_{im} = \sum_{k=1}^{K} \beta_k(X_k \times Engaged) + \sum_{k=1}^{K} \beta_k(X_k \times Random) \tag{7}$$

where $Random_{im}$ and $Engaged_{im}$ are indicator variables equal to 1 if the respondent belongs to the random or the engaged sample, respectively. X_k is a set of the following attributes and interaction terms: $Acres_{im}$, $PartVisibility_{im}$, $FullVisibility_{im}$, $Setback_{im}$, $Probability_{im}$, $Cost_{im}$, $Farm_i \times ASC_{im}$, $Forest_i \times ASC_{im}$, $Brownfield_i \times ASC_{im}$, and $Commercial_i \times ASC_{im}$. $Acres_{im}$ refers to the size of the installation (in acres), $PartVisibility_{im}$ and $FullVisibility_{im}$ are dummy variables equal to 1 if the installation is partly visible or completely visible, respectively, $Setback_{im}$ is to the setback distance (in 00's feet), $Probability_{im}$ is the likelihood of future solar development on farm and forest land, $Cost_{im}$ is the change in respondents' monthly electricity bill, and ASC_{im} is the status-quo alternative-specific constant which is an indicator variable equal to 1 for the status-quo choice and equal to 0 for either of the solar development options (Choices A and B). $Farm_i$, $Forest_i$, $Brownfield_i$, and $Commercial_i$ are all dummy variables equal to 1 if the choice set is framed around the respective land use.

In Equation (7), each solar attribute k is associated with a preference coefficient β_{ks}, where s represents whether the respondent belongs to the random or the engaged sample. A positive sign on β_{ks} indicates that respondents in sample s prefer the associated attribute associated with it, while a negative sign signals a dislike for it. The interaction of the ASC_{im} term with a land use type l and sample s represents the respective sample's preferences for developing solar arrays on the land use type l. If the coefficient associated with the triple interaction between land use type l, the ASC_{im} term, and sample$s$ is positive, it implies that the respondents in that sample prefer the status-quo option over the other alternatives and that they dislike having solar arrays on the associated land parcel. All coefficients are calculated using maximum likelihood procedures.

We obtain each sample's MWTP for attribute k by dividing the coefficient of that attribute with the negative of the cost coefficient associated with the respective group, β_{Costs}:

$$MWTP_k = -\frac{\beta_{ks}}{\beta_{Costs}} \tag{8}$$

Sample s's maximum WTP (or compensating variation) for a specified plan can be estimated by finding the price that makes the utility derived by the respondents in that sample from that particular plan, denoted as V_i^{1s}, equal to the utility they derive from the status quo option, denoted $V_i^{SQs\_l}$. Therefore, the compensating variation (CV) for each land use type l can be expressed as:

$$CV_l = \frac{V^{SQs\_l} - V^{1s*}}{\beta_{Costs}} \tag{9}$$

where V^{1s*} is the utility of the non-price attributes associated with the solar development plan under consideration, for sample s. Subtracting the CV associated with one land use type from another gives us the premium the average respondent would pay for switching solar panels from one to the other.

4 Data

4.1 Survey implementation

For the engaged sample, we recruited participants through an online invitation to attend a webinar on August 14, 2020 titled "Valuing Siting Options for Solar Energy in Rhode Island." The webinar was advertised by the University of Rhode Island's Cooperative Extension program on social media (Facebook and Instagram) and through email blasts using the Cooperative Extension email database and the Rhode Island Office of Energy Resources (RIOER) database. The webinar introduced the attendees to our study, along with its aims and objectives. The attendees were also informed that they would be receiving an email invitation to take our survey online the following day.

The survey was disseminated to the engaged sample in two rounds. On August 15, 2020, 104 individuals that had registered for our webinar were emailed a one-time, non-shareable link to our online survey. The link was emailed to all registrants, and not just to those who had attended. One week later, a reminder email was sent out to those registrants who had not responded.

For the random sample, we used best practices as outlined by Dillman et al. (2014) to design a mixed-mode, web-push survey (Lang et al., 2021). The mixed-mode aspect allows us to collect data both online (using Qualtrics) and on paper.

We used the 2019 Rhode Island voter registration database (publicly available from the Secretary of State) to draw a random sample of 3,000 individuals residing in RI. These data include name, address, age, party affiliation, and whether the individual participated in the last eight elections held. We adjusted sample selection probabilities to increase the odds of selecting younger people, those living in rural areas, and Republicans. We oversampled younger residents because we anticipated lower response rates from them. We also oversampled rural residents because this subgroup is more likely to be impacted by solar siting decisions. Finally, we elected to oversample Republicans because they comprise a smaller subset of the population in Rhode Island and have been found to be less responsive to surveys (Best and Krueger, 2012; Pearson-Merkowitz and Lang, 2020).

The survey was disseminated to the random sample in three rounds. The first round was mailed on September 4, 2020. Each envelope included an introductory letter that provided a link and unique access code to the online survey, and a $2 cash incentive. Two weeks later, non-respondents were sent a follow-up postcard as a reminder, which also gave the link and access code. In the third and final round (mailed two weeks after the second round), subjects who had not responded to either of the first two rounds of mailings were sent a paper survey.

Of the 104 webinar registrants, 48 attended the webinar, 30 of which responded to our survey. 20 non-attendee registrants also took the survey, giving us a total response rate of 48% for the engaged sample. Of the 2,796 surveys that were delivered to the random sample, we received 669 total responses (24% response rate). 510 were collected from the online mode and 159 were received by mail. We dropped 18 individuals who did not answer any of the choice questions (13 from the random sample and 5 from the engaged sample). Our final, pooled sample consists of 12,534 survey choices made by 701 individuals.

4.2 Summary statistics

To make our random sample representative of the RI population, we use state voter registration data to construct sample weights. Three key demographic variables also inform the construction of the weights: age, political affiliation, and rural/urban residence.[23] All engaged sample respondents were given a weight equal to 1, which is the average weight given to respondents in our weighted random sample.

Summary statistics are presented in Table 1. Columns 1 and 2 report the means and standard deviations for the (weighted) random and engaged samples, respectively. Column 3 presents the difference in means between the random and engaged samples, along with standard errors. Most characteristics are well-balanced between the groups. The largest differences lie in educational attainment and political affiliation. Over 90% of the engaged sample respondents have a college degree or higher, compared to 69% from the random sample. This is in line with Whitehead (1991) who finds that the interest group members have a higher level of education compared to the general sample. Only 4% of our engaged sample respondents are Republicans, compared to 14% in the random sample. On average, the engaged sample has fewer children living at home and a lower electricity bill, with the differences being statistically significant

[23] Table A1 in the online appendix reports the demographic distribution for these three variables in our unweighted, non-engaged sample, the population, and the weighted, non-engaged sample.

at the 5% and 10% level, respectively.

The two samples also differ in their attitudes towards different energy sources. While both groups have a positive attitude towards renewable energy sources (solar, wind, and hydro), the engaged sample displayed an extremely positive attitude. Every respondent in the engaged sample had a positive or very positive attitude towards solar, and their attitude towards onshore wind was slightly more positive compared to the non-engaged sample. The engaged sample also displayed a negative attitude towards natural gas, compared to the non-engaged sample, which was more neutral. Though both groups dislike nuclear energy sources, the engaged respondents are slightly more supportive of it on average. Both groups strongly dislike energy production from coal.

5 Results

5.1 Main results

Table 2 presents results obtained from logistic regressions estimating Equation (8). Column 1 reports coefficients from the CL model. Column 2 shows coefficients derived from estimating the HCL model, along with scale parameters associated with farm, forest, and commercial land use types.[24] Column 3 reports coefficients from the RPL model, while Column 4 presents standard deviations from the same model.

Results are broadly consistent across columns. We find that the coefficient on $Acres \times Random$ is positive and significant at the 1% level in all models, which demonstrates that respondents in the random sample prefer large solar installations. The coefficient on $Acres \times Engaged$ is also positive across the board, suggesting that engaged respondents also like large installations, though the coefficient it is only significant (at the 1% level) in the HCL model. This is likely because the small sample size of the engaged respondents precludes us from identifying effects that are small in magnitude. In addition, Figure A1 in the online appendix illustrates that installation size matters more to the engaged respondents than to the respondents in the random

[24] The indicator variable $Engaged$ is excluded from the list of scale parameters because it is not possible to include it along with interactions for each attribute. We also estimate alternative models (results not presented) that do not use the full set of attribute interactions but include $Engaged$ in the scale parameter, and do not find statistically different scale values for the engaged sample compared to the nonengaged sample.

sample. A greater percentage of engaged respondents reported that installation size was of “a lot” or “some” importance, and fewer reported that it was “not at all” important compared to non-engaged respondents.

The interaction coefficients $PartVisibility \times Random$ and $PartVisibility \times Engaged$ display a negative sign, though it is not statistically significant in any of the models. In contrast, the interactions of both groups with $FullVisibility$ are negative and statistically significant at the 1% level across all specifications. This implies that fully visible installations elicit a more negative response compared to partly visible installations from engaged and randomly selected respondents alike.

The coefficient on $Setback$ is positive for both groups and across all models, though there is some variation in significance. $Setback \times Random$ is weakly significant (at the 10% level) in the RPL model and insignificant in the CL and HCL models, and small in magnitude throughout. In contrast, the coefficient on $Setback \times Engaged$ is larger in magnitude, significant in the CL (at the 10% level) and HCL (at the 5% level) models, and insignificant in the RCL model. This provides suggestive evidence that both groups of respondents are slightly affected by setback distance. Additionally, the standard deviation in the RCL model associated with the attribute is significant for the random sample but not for the engaged sample, suggesting that preferences for setback distance are more homogeneous in the engaged sample.

$Probability \times Random$ is negative and significant at the 1% level across all models, which implies that as the probability of future residential development on farms and forests increases, randomly selected respondents are less likely to choose the status quo option. Contrastingly, engaged respondents are unaffected by the probability of development in the future, as suggested by the insignificant coefficient on $Probability \times Engaged$ in all models. This can also be seen in Figure A1, which depicts that almost 40% of the engaged sample respondents reported that they were “not at all” affected by future probability of development. The coefficients on $Cost \times Random$ and $Cost \times Engaged$ are consistently negative and significant at the 1% level, signifying that both groups dislike having a higher electric bill, which is a standard finding. More interesting, however, is that the cost coefficients associated with the engaged sample are lower in magnitude and statistically significantly different (at the 10% level) compared to the random sample, implying that the engaged respondents are less affected by higher costs. This can also be seen in Figure A1 where almost 36% of engaged respondents report $Cost$ to be "not at all" important, compared to only 18% of the respondents in the random sample.

We find that both groups of respondents have almost similar preferences regarding the land type on which solar is developed. The interaction term $Farm \times ASC$ is positive in all models, and for both the engaged and random samples, indicating respondents' dislike for having farmlands converted for solar development. However, the significance is not consistent. The coefficient is significant at the 1% level in the CL and HCL models and at the 10% level in the RPL model for the random sample. For the engaged sample, it is weakly significant (at the 10% level) only in the HCL model. The coefficients on $Forest \times ASC$ are positive and strongly significant across all models and for both groups. This indicates that both random and engaged respondents dislike solar development on forest lands. The coefficients on $Brownfield \times ASC \times Random$are negative and significant (at the 1% level) throughout, signifying the random sample respondents' preference for siting solar installations on brownfields. The $Brownfield \times ASC \times Engaged$ coefficients are also negative, though there are inconsistencies in significance. It is significant at the 1% level in the CL model and at the 5% level in the HCL model, but insignificant in the RPL model. Overall, it provides some evidence suggesting that engaged sample respondents also like having solar installations built on brownfield lands. The coefficients on $Commercial \times ASC$ are negative and significant in all models and for both groups, indicating that commercial sites are strongly and universally preferred for solar development.

Panel A of Table 3 reports MWTP estimates for all solar attributes for both groups, along with standard errors derived using the delta method. We find that randomly selected respondents have a monthly WTP between \$0.24 and \$0.28 for each additional acre of land to be developed for solar. The engaged sample respondents are willing to pay almost twice that, with WTP values ranging between \$0.40 and \$0.58 per month. The corresponding monthly WTP values for a 30 acre installation range between \$7.20 and \$8.40 for the random sample and between \$12 and \$17.40 for the engaged sample respondents. This is consistent with the engaged sample respondents' highly positive attitudes towards solar energy.

The MWTP for a partly visible installation is negative, insignificant, and small in magnitude for respondents in the random sample. The engaged sample respondents are willing to pay five to six times more to avoid visual disamenities from partly visible installations, though the value is only weakly significant (at the 10% level) in the RPL model. We find that both groups have a higher negative monthly MWTP for a fully visible installation compared to a partly visible one, and all estimates are significant at the 1% level. Additionally, the engaged sample's values are higher by a factor of almost three. On average, randomly sample respondents are willing to pay between \$6.21 and \$8.43 per month and engaged respondents between \$22.12 and \$27.28 per month to avoid seeing solar installations that are completely visible. This difference in WTP values is statistically significant at the 10% level in the CL and RPL models and at the 5% level in the HCL model.

The MWTP for $Setback$ is small for the random sample, ranging between \$0.98 and \$1.19 per month, and significant in the RPL model only. For the engaged sample, the MWTP for additional setback distance is larger, and ranges between \$4.05 and \$5.83 per month, suggesting that they care much more about setback than respondents in the random sample. The values for the $Probability$ attribute are negative and significant for the random sample, indicating that when the probability of future residential development increases, they require less compensation for the conversion of farms and forests into solar installations. The corresponding estimates for the engaged sample respondents are insignificant across all models, and very small in magnitude in the CL and HCL models, suggesting that the engaged respondents are not as affected as the random sample respondents by future residential development on farms and in forests.

In Panel B of Table 3, we present CV estimates for solar development on different land types. We make the assumption of a 10 acre solar installation (median size of solar installation in RI) that is completely visible, has a setback distance of 150 feet with a 0% probability of future residential development. Our results provide suggestive evidence that both engaged and random respondents dislike having solar installations on farmlands. On average, random sample respondents are willing to pay between \$12.68 and \$23.44 per month to avoid solar installations on farms. Engaged sample respondents are willing to pay between \$17.40 and \$36.53 for the same. The monthly WTP to avoid solar development on forests is large for both groups of respondents, and the estimate is significant at the 1% level across models. We find that random sample respondents are willing to pay between \$40.58 and \$49.49 per month for avoiding solar development on forested land. Engaged sample respondents are willing to pay even higher amounts, between \$82.08 and \$198.50 per month, indicating a very strong aversion to having forests converted into solar installations.

Both groups support developing commercial sites into solar, with respondents in the random sample willing to pay between \$14.31 and \$20.72 per month and engaged respondents about three times as much – between \$43.71 and \$73.88 per month to have solar installations built on commercial land. We also find positive WTP values for $Brownfield$ for the random sample, indicating that they prefer having solar installations on brownfields with a monthly WTP between \$9.97 and \$15.07. For the engaged group, the WTP value for solar development on brownfields is larger and positive, ranging between \$18.90 and \$57.73, although it is significant in the CL model only. The large differences in the WTP and CV magnitudes between the engaged and random samples could partly be driven by the small cost coefficients associated with the engaged respondents. Since the engaged sample has larger coefficients associated with most solar attributes (compared to the non-engaged sample), the combination of a large numerator and a small denominator leads to very small WTP and CV values.

5.2 Robustness check with online-only responses

We conduct a robustness analysis by dropping 148 respondents from the random sample who had responded by mail. This eliminates any variation in responses that could be caused due to differences in survey mode. We generate new survey weights for the smaller random sample of 508 online respondents and assign a weight equal to 1 to the engaged respondents.[25] The final sample consists of 9,933 choices made by 553 respondents.

Panel A of Table 4 reports the MWTP values for all solar siting attributes for both samples.[26] Overall, the results from the online-only sample are qualitatively similar to our main results. Respondents from both samples like larger installations, but engaged respondents are willing to pay almost two times more than random sample respondents for each additional acre of solar development. Assuming a 30 acre installation, random sample respondents have a monthly MWTP between \$8.10 and \$9.60, while engaged respondents have a monthly MWTP between \$12 and \$18. Negative MWTP values for $PartVisibility$ indicate that both sets of respondents dislike partly visible installations, though the coefficient is significant only for the random sample in the RPL model. In line with our main results, engaged respondents are willing to pay two to four times more per month than random sample respondents to avoid seeing fully visible installations. This value ranges between -\$6.52 and -\$8.40 per month for random sample respondents and between -\$20.72 and -\$27.28 for engaged respondents. Engaged respondents care more about setback distance compared to random sample respondents and are willing to pay between \$4.92 and \$5.86 per month for increased setback distance. $Probability \times Random$ is negative and significant in all models at the 1% level, indicating that random sample respondents are willing to pay between \$5 and \$12 per month for a 25% reduction in the probability of future residential development. However, $Probability \times Engaged$ is insignificant throughout, which implies that engaged respondents are unaffected by probability of residential development on farms and forest lands, a finding that is similar to our main results.

We calculate CV values for solar development on various land types and present the estimates in Panel B of Table 4. As with our main results, we assume a 10 acre, fully visible installation, with a 150 feet setback distance and a 0% probability of future residential development. Our results are consistent with the results of the main

[25] Survey weights are generated using the same procedure that is described in Section 4.2.
[26] The logistic regression coefficients that are used to estimate CV values in Table 4 are presented in Table A1 of the online appendix.

model. Respondents from both samples dislike solar development on farmlands, with the random sample respondents willing to pay between \$17.11 and \$25.13 per month and the engaged respondents willing to pay around \$36 to avoid converting farms into solar installations. We find that respondents from both samples have a high WTP to avoid solar development on forested lands, but engaged respondents are willing to pay two to three times more than random sample respondents. The upper bound in the difference between WTP values is slightly less compared to our main results, where engaged respondents are willing to pay up to four times more to avoid solar development on forested lands. On average, the monthly WTP to avoid solar development in forests ranges between \$42.59 and \$54.29 for random sample respondents and between \$82.08 and \$157.98 for the engaged respondents.

We find positive WTP values for solar development on $Commercial$ for respondents from both groups, with random sample respondents willing to pay between \$12.75 and \$19.90 per month and engaged respondents between \$40.80 and \$196.30 per month (three times higher in the CL and HCL models and ten times higher in the RPL model) to have solar installations built on commercial land. This difference in WTP values is much higher than in our main results, where engaged respondents are willing to pay only three times as much as random respondents for solar development on commercial sites. Our results indicate that both random sample respondents and engaged respondents support solar development on brownfields. On average, respondents from the random sample are willing to pay between \$9.82 and \$15.45 for building solar installations on brownfields. As with our main results, engaged respondents are willing to pay almost two times more for the same, between \$18.50 and \$38.04 per month, though the estimate is significant (at the 10% level) in the CL and RPL models only.

6 Conclusion

This paper analyzes differences between a random sample and a conveniently obtained sample of engaged respondents. Engaged respondents were recruited from a registration list of 104 individuals that had signed up for a webinar on “Valuing Siting Options for Solar Energy in RI”. The random sample is comprised of respondents to a survey that was distributed to 2,794 randomly selected RI residents. Our final sample includes 45 engaged and 656 randomly selected respondents. We elicit residents' preferences for the following solar energy siting attributes: installation size, visibility, setback distance, probability of future residential development, change in electricity bill, and current land use of the proposed solar site. WTP values are estimated using the CE design and logistic regression models.

Our results indicate that while both groups of respondents have similar preferences, there are differences in their MWTP for several attributes. Both engaged and random sample respondents prefer larger installations, but the engaged sample is willing to pay between $12 and $17.40 per month for a 30 acre installation, which is two times more than what the respondents in the random sample are willing to pay. Engaged respondents' monthly MWTP to avoid visual disamenities from fully visible installations is between $22.12 and $27.28, which is three times more than the random sample respondents' values (between $6.21 and $8.43). Respondents in the random sample are mainly unaffected by setback distance, but the engaged respondents care much more about it and are willing to pay between $4.05 and $5.83 per month for increased setback distance. Random sample respondents are less likely to choose the status quo option when the probability of future residential development on farms and forests increases, but engaged respondents are not affected by it.

We obtain total WTP values for solar development on various land types by assuming a 10 acre, fully visible installation with a setback distance of 150 feet and 0% probability of future residential development. Both groups equally dislike having solar installations on farmlands, with a total WTP between $12.68 and $23.44 per month for the random sample and between $17.40 and $36.53 for the engaged sample.

We find large, negative WTP values for solar development on forests for both groups, but the magnitude is two to four times larger for the engaged sample. On average, random sample respondents are willing to pay between $40.58 and $49.49 per month, and engaged respondents are willing to pay between $82.08 and $198.50 per month to avoid solar development on forests. Random sample respondents support developing commercial land into solar with a total WTP value ranging between $14.31 and $20.72 per month, and engaged respondents are willing to pay between $43.71

and \$73.88 per month for the same, which is over three times more than random sample respondents' WTP. Respondents in the random sample also like solar development on brownfield sites, with a monthly total WTP between \$9.97 and \$15.07. Engaged respondents similarly support the conversion of brownfield lands into solar, with a total WTP between \$18.90 and \$57.73.

In conclusion, we find that the engaged respondents' WTP for a majority of solar siting attributes is several times larger compared to random sample respondents. However, the direction of preferences is the same for all attributes, indicating that both groups have similar likes and dislikes overall. Thus, while caution must be exercised when considering the WTP values of an engaged group of individuals for policy purposes, inferences about overall preferences can be considered to be fairly representative of the general population.

Tables

Table 1: Summary statistics

Variable	Random sample (std. dev.)	Engaged sample (std. dev.)	Difference in means (std. error)
Household income (000's)	106.75	99.39	-7.36
	(51.42)	(46.58)	(7.60)
College education (1 = yes)	0.69	0.91	0.23***
	(0.46)	(0.91)	(0.05)
Children at home (1 = yes)	0.35	0.21	-0.14**
	(0.48)	(0.41)	(0.07)
Female (1 = yes)	0.53	0.44	-0.08
	(0.50)	(0.50)	(0.08)
Age	50.81	53.02	2.21
	(15.31)	(16.45)	(2.56)
Homeowner (1 = yes)	0.79	0.78	-0.01
	(0.41)	(0.42)	(0.07)
Years of residence	15.04	14.50	-0.54
	(6.78)	(7.29)	(1.12)
Employed (1 = yes)	0.67	0.67	0.00
	(0.47)	(0.48)	(0.07)
Electricity bill ($/month)	120.74	105.56	-15.19*
	(54.50)	(54.93)	(8.47)
Solar panels at home (1 = yes)	0.05	0.11	0.06
	(0.22)	(0.32)	(0.05)
Urban (1 = yes)	0.23	0.20	-0.03
	(0.42)	(0.40)	(0.06)
Suburban (1 = yes)	0.63	0.56	-0.08
	(0.48)	(0.50)	(0.08)
Democrat (1 = yes)	0.38	0.40	0.02
	(0.49)	(0.50)	(0.08)
Republican (1 = yes)	0.14	0.04	-0.10***
	(0.35)	(0.21)	(0.03)
Energy attitudes (1 = positive)			
Solar	0.88	1.00	0.12***
	(0.33)	(0.00)	(0.01)
Offshore wind	0.81	0.87	0.05
	(0.39)	(0.34)	(0.05)
Onshore wind	0.76	0.89	0.13**
	(0.43)	(0.32)	(0.05)
Hydro	0.68	0.78	0.10
	(0.47)	(0.42)	(0.07)
Natural gas	0.49	0.29	-0.20***
	(0.50)	(0.46)	(0.07)
Nuclear	0.24	0.38	0.14*
	(0.43)	(0.49)	(0.07)
Coal	0.07	0.00	-0.07***
	(0.25)	(0.00)	(0.01)

Note: All data come from survey responses. Household income and electricity bill values come from a multiple choice question that included several ranges. We assign people the middle value of their chosen range.

Table 2: Attribute coefficients from logit regressions

Variable	Conditional Logit	Heteroscedastic logit	Random Parameters Logit Mean	SD
Acres × Random	0.010***	0.019***	0.016***	0.042***
	(0.002)	(0.004)	(0.003)	(0.005)
Acres × Engaged	0.011	0.026***	0.024	0.046***
	(0.007)	(0.009)	(0.015)	(0.016)
Part visibility × Random	-0.066	-0.067	-0.132	0.202
	(0.061)	(0.086)	(0.088)	(0.484)
Part visibility × Engaged	-0.294	-0.334	-0.468	0.506
	(0.198)	(0.262)	(0.299)	(0.365)
Full visibility × Random	-0.313***	-0.408***	-0.524***	0.750***
	(0.074)	(0.103)	(0.109)	(0.203)
Full visibility × Engaged	-0.783***	-1.049***	-1.084***	0.382
	(0.233)	(0.306)	(0.376)	(0.425)
Setback (00's ft) × Random	0.042	0.066	0.077*	0.298***
	(0.031)	(0.043)	(0.044)	(0.100)
Setback (00's ft) × Engaged	0.151*	0.262**	0.198	0.102
	(0.089)	(0.126)	(0.145)	(0.152)
Probability × Random	-0.008***	-0.014***	-0.031***	0.087***
	(0.002)	(0.004)	(0.008)	(0.014)
Probability × Engaged	-0.002	-0.003	-0.024	0.150***
	(0.008)	(0.014)	(0.032)	(0.058)
Cost × Random	-0.043***	-0.066***	-0.065***	
	(0.002)	(0.007)	(0.004)	
Cost × Engaged	-0.029***	-0.045***	-0.049***	
	(0.008)	(0.012)	(0.013)	
Land use interactions				
Farm × ASC × Random	0.822***	1.425***	0.571*	3.688***
	(0.135)	(0.268)	(0.331)	(0.582)
Farm × ASC × Engaged	0.566	1.248*	0.284	0.870
	(0.420)	(0.641)	(0.642)	(0.981)
Forest × ASC × Random	1.596***	3.071***	2.946***	4.058***
	(0.134)	(0.473)	(0.409)	(0.752)
Forest × ASC × Engaged	1.913***	4.116***	9.175**	11.622
	(0.471)	(0.971)	(3.634)	(7.947)
Brownfield × ASC × Random	-0.793***	-0.780***	-1.203***	0.114
	(0.128)	(0.149)	(0.166)	(0.229)
Brownfield × ASC × Engaged	-1.534***	-1.244**	-3.380	2.064
	(0.565)	(0.601)	(6.735)	(5.886)
Commercial × ASC × Random	-1.035***	-1.067***	-1.494***	0.089
	(0.132)	(0.189)	(0.171)	(0.275)
Commercial × ASC × Engaged	-2.562***	-2.360***	-3.266***	0.409
	(0.769)	(0.883)	(0.989)	(0.968)
Heteroskedastic variables				
Farm		-0.454***		
		(0.120)		
Forest		-0.741***		
		(0.141)		
Commercial		-0.027		
		(0.135)		
Observations	12,534	12,534	12,534	
AIC	7879.706	7832.115	7064.510	
BIC	8028.430	8003.148	7347.086	

Note: Acres refers to the size of the solar installation in acres. Part visibility and Full visibility are dummy variables = 1 if a solar installation is partially or completely visible, respectively. ASC is the status-quo alternative-specific constant, or a dummy variable = 1 for the status-quo choice and 0 otherwise. Cost is in terms of USD per person per month. Cluster robust standard errors are in parentheses. *,**,*** indicate significance at the 90%, 95% and 99% level, respectively.

Table 3: Willingness to pay estimates for solar attributes

Attribute	Conditional logit	Heteroskedastic logit	Random parameters logit
Panel A: Marginal WTP			
Acres × Random	$0.24***	$0.28***	$0.25***
	(0.05)	(0.05)	(0.05)
Acres × Engaged	$0.40	$0.58**	$0.48*
	(0.27)	(0.24)	(0.27)
Part visibility × Random	-$1.54	-$1.02	-$2.04
	(1.43)	(1.32)	(1.36)
Part visibility × Engaged	-$10.26	-$7.43	-$9.56*
	(6.45)	(5.57)	(5.51)
Full visibility × Random	-$7.30***	-$6.18***	-$8.11***
	(1.79)	(1.54)	(1.67)
Full visibility × Engaged	-$27.28***	-$23.34***	-$22.12***
	(9.81)	(7.82)	(7.95)
Setback (00's ft) × Random	$0.98	$1.01	$1.19*
	(0.74)	(0.64)	(0.69)
Setback (00's ft) × Engaged	$5.25*	$5.83**	$4.05
	(3.10)	(2.78)	(2.80)
Probability × Random	-$0.19***	-$0.22***	-$0.48***
	(0.05)	(0.06)	(0.12)
Probability × Engaged	-$0.08	-$0.06	-$0.50
	(0.29)	(0.32)	(0.65)
Panel B: Total WTP			
Farm			
Random sample	-$22.54***	-$23.44***	-$12.68**
	(3.08)	(3.13)	(5.05)
Engaged sample	-$35.13**	-$36.53**	-$17.04
	(15.92)	(16.34)	(11.20)
Forest			
Random sample	-$40.58***	-$48.38***	-$49.49***
	(3.29)	(4.80)	(6.12)
Engaged sample	-$82.08***	-$100.31***	-$198.50***
	(22.34)	(27.65)	(72.84)
Commercial			
Random sample	$20.72***	$14.31***	$19.30***
	(3.13)	(2.36)	(2.75)
Engaged sample	$73.88**	$43.71**	$55.43***
	(29.98)	(18.66)	(21.47)
Brownfield			
Random sample	$15.07***	$9.97***	$14.80***
	(2.93)	(2.30)	(2.64)
Engaged sample	$38.04*	$18.90	$57.73
	(20.04)	(12.16)	(129.91)

Notes: Welfare estimates are in USD per household per month. Estimates in Panel A represent marginal WTP values. In Panel B, the estimates represent total WTP values and assume a 10 acre, fully visible installation with a setback of 150 feet, and a 0% probability of development in the future. In both panels, standard errors are calculated using the delta method and are in parentheses. *,**,*** indicate significance at the 90%, 95% and 99% level, respectively.

Table 4: Willingness to pay estimates for solar attributes for online respondents

Attribute	Conditional logit	Heteroskedastic logit	Random parameters logit
Panel A: Marginal WTP			
Acres × Random	$0.27***	$0.32***	$0.28***
	(0.06)	(0.06)	(0.06)
Acres × Engaged	$0.40	$0.60**	$0.58**
	(0.27)	(0.24)	(0.27)
Part visibility × Random	-$2.36	-$1.70	-$2.93**
	(1.66)	(1.52)	(1.45)
Part visibility × Engaged	-$10.26	-$7.39	-$7.82
	(6.46)	(5.55)	(5.01)
Full visibility × Random	-$7.78***	-$6.52***	-$8.40***
	(1.91)	(1.60)	(1.86)
Full visibility × Engaged	-$27.28***	-$23.09***	-$20.72***
	(9.82)	(7.66)	(6.74)
Setback (00's ft) × Random	$1.41*	$1.15	$1.03
	(0.85)	(0.73)	(0.79)
Setback (00's ft) × Engaged	$5.25*	$5.86**	$4.92*
	(3.10)	(2.76)	(2.56)
Probability × Random	-$0.20***	-$0.23***	-$0.48***
	(0.06)	(0.08)	(0.12)
Probability × Engaged	-$0.08	-$0.05	-$0.96
	(0.29)	(0.33)	(1.00)
Panel B: Total WTP			
Farm			
Random sample	-$23.90***	-$25.13***	-$17.11***
	(3.51)	(3.57)	(5.00)
Engaged sample	-$35.13**	-$36.53**	-$1.43
	(15.92)	(16.33)	(24.13)
Forest			
Random sample	-$42.59***	-$52.50***	-$54.29***
	(3.81)	(6.02)	(6.65)
Engaged sample	-$82.08***	-$104.22***	-$157.98***
	(22.34)	(29.04)	(53.17)
Commercial			
Random sample	$19.90***	$12.75***	$18.91***
	(3.61)	(2.54)	(3.23)
Engaged sample	$73.88**	$40.80**	$196.30***
	(29.98)	(17.63)	(75.03)
Brownfield			
Random sample	$15.45***	$9.82***	$15.45***
	(3.39)	(2.62)	(3.00)
Engaged sample	$38.04*	$18.50	$33.15*
	(20.04)	(12.14)	(19.76)

Notes: Welfare estimates are in USD per household per month. Estimates in Panel A represent marginal WTP values. In Panel B, the estimates represent total WTP values and assume a 10 acre, fully visible installation with a setback of 150 feet, and a 0% probability of development in the future. In both panels, standard errors are calculated using the delta method and are in parentheses. *,**,*** indicate significance at the 90%, 95% and 99% level, respectively.

Appendix

This appendix provides supplemental figures and tables to our main results.

Figure A1 represents how important each attribute was for the respondents in the random and engaged samples while making decisions. Land use is the most important attribute for respondents in both samples. $Size$ is more important to engaged respondents than random sample respondents. Respondents in the random sample are slightly more affected by visibility than engaged respondents. The importance of $Setback$ is equally distributed between the two samples. Random sample respondents care more about future probability of residential development than engaged respondents, who find $Probability$ to be the least important attribute. $Cost$ is the second most important attribute for random sample respondents but for the engaged respondents it is the second least important, a finding that plays out in the main results as well.

Table A1 presents CL, HCL, and RPL coefficients obtained from estimating Equation (8) on the sample of online respondents. Overall, the results from this subsample are qualitatively similar to our main results. Both engaged and random sample respondents like larger installations, as indicated by the positive coefficients on $Acres \times Random$ and $Acres \times Engaged$. Both sets of respondents also dislike partly visible installations, though the coefficient is significant only for the random sample in the RPL model. The coefficients on $FullVisibility \times Random$ and $FullVisibility \times Engaged$ are negative and significant (at the 1% level) across all models, suggesting that fully visible installations are disliked by both groups of respondents. $Setback \times Random$ is significant (at the 10% level) only in the CL model, but the coefficients on $Setback \times Engaged$ are significant across all models, suggesting that engaged respondents are more responsive to setback distance compared to random sample respondents. $Probability \times Random$ is negative and significant in all models at the 1% level, indicating that random sample respondents are less likely to select the status-quo option as the future probability of residential development increases. However, $Probability \times Engaged$ is insignificant throughout, which implies that engaged respondents are unaffected by future residential development on farms and forest lands. Similar to our main results, respondents from both samples dislike having a higher electric bill (as indicated by the negative and significant coefficients on $Cost \times Random$ and $Cost \times Engaged$), though the smaller magnitude associated with the engaged respondents indicates that they are less affected by higher costs.

Figure A1: Importance of attributes while making choices by sample groups

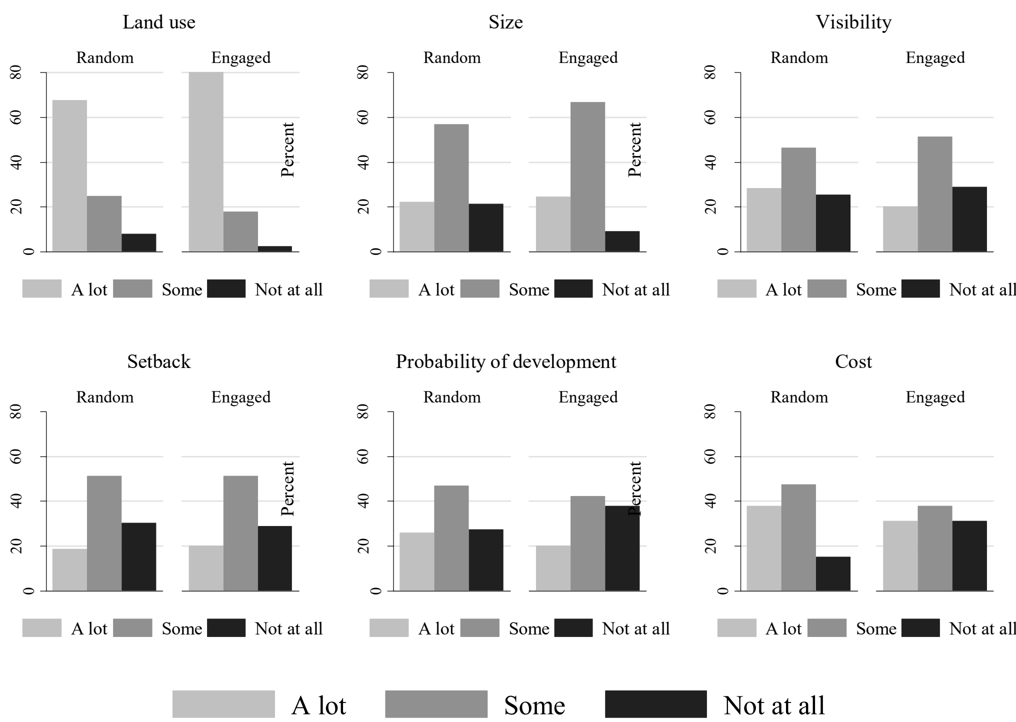


Table A1: Attribute coefficients from logit regressions for online respondents

Variable	Conditional Logit	Heteroscedastic logit	Random Parameters Logit Mean	SD
Acres × Random	0.012***	0.021***	0.018***	0.045***
	(0.003)	(0.004)	(0.004)	(0.005)
Acres × Engaged	0.011	0.027***	0.027**	0.036***
	(0.007)	(0.009)	(0.013)	(0.011)
Part visibility × Random	-0.101	-0.113	-0.190**	0.209
	(0.071)	(0.100)	(0.096)	(0.264)
Part visibility × Engaged	-0.294	-0.335	-0.368	0.601
	(0.198)	(0.262)	(0.254)	(0.423)
Full visibility × Random	-0.332***	-0.432***	-0.547***	0.442
	(0.079)	(0.110)	(0.124)	(0.291)
Full visibility × Engaged	-0.783***	-1.048***	-0.975***	0.057
	(0.233)	(0.309)	(0.275)	(0.303)
Setback (00's ft) × Random	0.060*	0.076	0.067	0.350***
	(0.036)	(0.049)	(0.051)	(0.107)
Setback (00's ft) × Engaged	0.151*	0.266**	0.232*	0.128
	(0.089)	(0.127)	(0.132)	(0.109)
Probability × Random	-0.009***	-0.016***	-0.031***	0.079***
	(0.003)	(0.005)	(0.008)	(0.014)
Probability × Engaged	-0.002	-0.002	-0.045	0.197**
	(0.008)	(0.015)	(0.046)	(0.087)
Cost × Random	-0.043***	-0.066***	-0.065***	
	(0.003)	(0.008)	(0.005)	
Cost × Engaged	-0.029***	-0.045***	-0.047***	
	(0.008)	(0.012)	(0.013)	
Land use interactions				
Farm × ASC × Random	0.893***	1.560***	0.849**	3.466***
	(0.151)	(0.315)	(0.333)	(0.537)
Farm × ASC × Engaged	0.566	1.279**	-0.289	4.421
	(0.420)	(0.649)	(1.083)	(2.913)
Forest × ASC × Random	1.690***	3.376***	3.268***	4.318***
	(0.155)	(0.599)	(0.464)	(0.713)
Forest × ASC × Engaged	1.913***	4.350***	7.081***	8.391***
	(0.471)	(1.069)	(2.433)	(3.252)
Brownfield × ASC × Random	-0.785***	-0.758***	-1.269***	0.034
	(0.146)	(0.169)	(0.186)	(0.211)
Brownfield × ASC × Engaged	-1.534***	-1.218**	-1.916**	1.378
	(0.565)	(0.599)	(0.835)	(1.489)
Commercial × ASC × Random	-0.975***	-0.953***	-1.495***	0.225
	(0.151)	(0.200)	(0.193)	(0.273)
Commercial × ASC × Engaged	-2.562***	-2.230***	-9.596***	5.641***
	(0.769)	(0.860)	(3.152)	(2.004)
Heteroskedastic variables				
Farm		-0.460***		
		(0.139)		
Forest		-0.806***		
		(0.166)		
Commercial		0.017		
		(0.151)		
Observations	9,933	9,933	9,933	
AIC	6230.385	6184.403	5542.923	
BIC	6374.458	6350.087	5816.661	

Note: Acres refers to the size of the solar installation in acres. Part visibility and Full visibility are dummy variables = 1 if a solar installation is partially or completely visible, respectively. ASC is the status-quo alternative-specific constant, or a dummy variable = 1 for the status-quo choice and 0 otherwise. Cost is in terms of USD per person per month. Cluster robust standard errors are in parentheses. *,**,*** indicate significance at the 90%, 95% and 99% level, respectively.

REFERENCES

Abashidze, N. (2019). *Essays on Economic and Health Effects of Land Use Externalities.* [North Carolina State University]. https://repository.lib.ncsu.edu/handle/1840.20/38420

Altonji, M., Lang, C., & Puggioni, G. (2016). Can urban areas help sustain the preservation of open space? Evidence from statewide referenda. *Ecological Economics*, *130*, 82–91. https://doi.org/10.1016/j.ecolecon.2016.06.026

Bates, A., & Firestone, J. (2015). A comparative assessment of proposed offshore wind power demonstration projects in the United States. *Energy Research & Social Science*, *10*, 192–205. https://doi.org/10.1016/j.erss.2015.07.007

Beatty, T. K. M., Lade, G. E., & Shimshack, J. (2021). Hurricanes and Gasoline Price Gouging. *Journal of the Association of Environmental and Resource Economists*, *8*(2), 347–374. https://doi.org/10.1086/712419

Best, S. J., & Krueger, B. S. (2012). *Exit polls: Surveying the American electorate, 1972-2010*. SAGE/CQ Press,.

Bishop, K. C., Kuminoff, N. V., Banzhaf, H. S., & Boyle, K. J. (2019). Best Practices in Using Hedonic Property Value Models for Welfare Measurement. *Review of Environmental Economics and Policy*, 43.

Boslett, A., Guilfoos, T., & Lang, C. (2019). Valuation of the External Costs of Unconventional Oil and Gas Development: The Critical Importance of Mineral Rights Ownership. *Journal of the Association of Environmental and Resource Economists*, *6*(3), 531–561. https://doi.org/10.1086/702540

Botelho, A., Lourenço-Gomes, L., Pinto, L., Sousa, S., & Valente, M. (2017). Accounting for local impacts of photovoltaic farms: The application of two stated preferences approaches to a case-study in Portugal. *Energy Policy*, *109*, 191–198. https://doi.org/10.1016/j.enpol.2017.06.065

Bowker, J. M., & Stoll, J. R. (1988). Use of Dichotomous Choice Nonmarket Methods to Value the Whooping Crane Resource. *American Journal of Agricultural Economics*, *70*(2), 372–381. https://doi.org/10.2307/1242078

Boyle, K. J., Boatwright, J., Brahma, S., & Xu, W. (2019). NIMBY, not, in siting community wind farms. *Resource and Energy Economics*, *57*, 85–100. https://doi.org/10.1016/j.reseneeco.2019.04.004

Brittan, G. G. (2001). Wind, energy, landscape: Reconciling nature and technology. *Philosophy & Geography*, *4*(2), 169–184. https://doi.org/10.1080/10903770124626

Carlisle, J. E., Kane, S. L., Solan, D., Bowman, M., & Joe, J. C. (2015). Public attitudes regarding large-scale solar energy development in the U.S. *Renewable and Sustainable Energy Reviews*, *48*, 835–847. https://doi.org/10.1016/j.rser.2015.04.047

Carlisle, J. E., Kane, S. L., Solan, D., & Joe, J. C. (2014). Support for solar energy: Examining sense of place and utility-scale development in California. *Energy Research & Social Science*, *3*, 124–130. https://doi.org/10.1016/j.erss.2014.07.006

Carr-Harris, A., & Lang, C. (2019). Sustainability and tourism: The effect of the United States' first offshore wind farm on the vacation rental market. *Resource*

and Energy Economics, *57*, 51–67. https://doi.org/10.1016/j.reseneeco.2019.04.003
Carson, R. T., & Groves, T. (2007). Incentive and informational properties of preference questions. *Environmental and Resource Economics*, *37*(1), 181–210. https://doi.org/10.1007/s10640-007-9124-5
Chen, Z., Swallow, S. K., & Yue, I. T. (2020). Non-participation and Heterogeneity in Stated: A Double Hurdle Latent Class Approach for Climate Change Adaptation Plans and Ecosystem Services. *Environmental and Resource Economics*, *77*(1), 35–67. https://doi.org/10.1007/s10640-020-00434-z
Czajkowski, M., Hanley, N., & LaRiviere, J. (2016). Controlling for the Effects of Information in a Public Goods Discrete Choice Model. *Environmental and Resource Economics*, *63*(3), 523–544. https://doi.org/10.1007/s10640-014-9847-z
Dastrup, S. R., Graff Zivin, J., Costa, D. L., & Kahn, M. E. (2012). Understanding the Solar Home price premium: Electricity generation and "Green" social status. *European Economic Review*, *56*(5), 961–973. https://doi.org/10.1016/j.euroecorev.2012.02.006
Davis, L. W. (2011). THE EFFECT OF POWER PLANTS ON LOCAL HOUSING VALUES AND RENTS. *The Review of Economics and Statistics*, *93*(4), 1391–1402.
Davis, M., Smith, C., White, B., Goldstein, R., Sun, X., Cox, M., Curtin, G., Manghani, R., Rumery, S., Silver, C., & Baca, J. (2021). *U.S. Solar Market Insight*. Wood Mackenzie and Solar Energy Industries Association.
de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, *110*(9), 2964–2996. https://doi.org/10.1257/aer.20181169
de Chaisemartin, C., D'Haultfœuille, X., & Guyonvarch, Y. (2019). Fuzzy differences-in-differences with Stata. *The Stata Journal*, *19*(2), 435–458. https://doi.org/10.1177/1536867X19854019
Denholm, P., & Margolis, R. M. (2008). Land-use requirements and the per-capita solar footprint for photovoltaic generation in the United States. *Energy Policy*, *36*(9), 3531–3543. https://doi.org/10.1016/j.enpol.2008.05.035
DeShazo, J. R., & Fermo, G. (2002). Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. *Journal of Environmental Economics and Management*, *44*(1), 123–143. https://doi.org/10.1006/jeem.2001.1199
Devine-Wright, P. (2009). Rethinking NIMBYism: The role of place attachment and place identity in explaining place-protective action. *Journal of Community & Applied Social Psychology*, *19*(6), 426–441. https://doi.org/10.1002/casp.1004
Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. John Wiley & Sons.
Dröes, M. I., & Koster, H. R. A. (2016). Renewable energy and negative externalities: The effect of wind turbines on house prices. *Journal of Urban Economics*, *96*, 121–141. https://doi.org/10.1016/j.jue.2016.09.001
Dröes, M. I., & Koster, H. R. A. (2021). Wind turbines, solar farms, and house prices. *Energy Policy*, *155*, 112327. https://doi.org/10.1016/j.enpol.2021.112327

EcoRI. (2020). 6 Acres of Solar Energy to Help Glocester Farm Pay Bills. *EcoRI News*. https://www.ecori.org/renewable-energy/2020/8/7/solar-energy-helps-farm-endure

Edwards, S. F., & Anderson, G. D. (1987). Overlooked Biases in Contingent Valuation Surveys: Some Considerations. *Land Economics*, *63*(2), 168–178. https://doi.org/10.2307/3146578

EIA - Annual Energy Outlook 2018. (2018, July 17). https://www.eia.gov/outlooks/aeo/

EIA Annual Energy Outlook 2021. (2021). U.S. Energy Information Administration. https://www.eia.gov/outlooks/aeo/

EIA International Energy Outlook. (2019). U.S. Energy Information Administration. https://www.eia.gov/outlooks/ieo/

Farhar, B. C. (1994). Trends in US Public Perceptions and Preferences on Energy and Environmental Policy. *Annual Review of Energy and the Environment*, *19*(1), 211–239. https://doi.org/10.1146/annurev.eg.19.110194.001235

Farhar, B. C., Hunter, L. M., Kirkland, T. M., & Tierney, K. J. (2010). *Community Response to Concentrating Solar Power in the San Luis Valley: October 9, 2008 - March 31, 2010* (NREL/SR-550-48041). National Renewable Energy Lab. (NREL), Golden, CO (United States). https://doi.org/10.2172/983406

Firestone, J., Bidwell, D., Gardner, M., & Knapp, L. (2018). Wind in the sails or choppy seas?: People-place relations, aesthetics and public support for the United States' first offshore wind project. *Energy Research & Social Science*, *40*, 232–243. https://doi.org/10.1016/j.erss.2018.02.017

Gaur, V., & Lang, C. (2020). *Property value impacts of Commercial-Scale Solar Energy in Massachusetts and Rhode Island*. https://web.uri.edu/coopext/valuing-siting-options-for-commercial-scale-solar-energy-in-rhode-island/

Geoghegan, J., Lynch, L., & Bucholtz, S. J. (2006). Are Agricultural Land Preservation Programs Self-financing? In *Economics and Contemporary Land Use Policy* (pp. 175–195). Routledge. https://doi.org/10.4324/9781936331659-17

Gibbons, S. (2015). Gone with the wind: Valuing the visual impacts of wind turbines through house prices. *Journal of Environmental Economics and Management*, *72*, 177–196. https://doi.org/10.1016/j.jeem.2015.04.006

Greenberg, M. (2009). Energy sources, public policy, and public preferences: Analysis of US national and site-specific data. *Energy Policy*, *37*(8), 3242–3249. https://doi.org/10.1016/j.enpol.2009.04.020

Gross, S. (2020). *Renewables, land use, and local opposition in the United States* (p. 24). The Brookings Institution. https://docs.wind-watch.org/FP_20200113_renewables_land_use_local_opposition_gross.pdf

Hausman, C., & Muehlenbachs, L. (2018). Price Regulation and Environmental Externalities: Evidence from Methane Leaks. *Journal of the Association of Environmental and Resource Economists*, *6*(1), 73–109. https://doi.org/10.1086/700301

Hedt, B. L., & Pagano, M. (2011). Health indicators: Eliminating bias from convenience sampling estimators. *Statistics in Medicine*, *30*(5), 560–568. https://doi.org/10.1002/sim.3920

Hensher, D., Louviere, J., & Swait, J. (1998). Combining sources of preference data. *Journal of Econometrics*, *89*(1), 197–221. https://doi.org/10.1016/S0304-4076(98)00061-X

Herriges, J., Kling, C., Liu, C.-C., & Tobias, J. (2010). What are the consequences of consequentiality? *Journal of Environmental Economics and Management*, *59*(1), 67–81. https://doi.org/10.1016/j.jeem.2009.03.004

Hoen, B., & Atkinson-Palombo, C. (2016). Wind Turbines, Amenities and Disamenitites: Astudy of Home Value Impacts in Densely Populated Massachusetts. *Journal of Real Estate Research*, *38*(4), 473–504. https://doi.org/10.5555/0896-5803-38.4.473

Hoen, B., Brown, J. P., Jackson, T., Thayer, M. A., Wiser, R., & Cappers, P. (2015). Spatial Hedonic Analysis of the Effects of US Wind Energy Facilities on Surrounding Property Values. *The Journal of Real Estate Finance and Economics*, *51*(1), 22–51. https://doi.org/10.1007/s11146-014-9477-9

Hoen, B., Firestone, J., Rand, J., Elliott, D., Hübner, G., Pohl, J., Wiser, R., Lantz, E., Haac, R., & Kaliski, K. (2019). Attitudes of U.S. Wind Turbine Neighbors: Analysis of a Nationwide Survey. *Energy Policy*, *134*. https://doi.org/10.1016/j.enpol.2019.110981

Hoen, B., Wiser, R., Thayer, M., & Cappers, P. (2012). Residential Photovoltaic Energy Systems in California: The Effect on Home Sales Prices. *Contemporary Economic Policy*, *31*(4), 708–718. https://doi.org/10.1111/j.1465-7287.2012.00340.x

Hultsch, D. F., MacDonald, S. W. S., Hunter, M. A., Maitland, S. B., & Dixon, R. A. (2002). Sampling and generalisability in developmental research: Comparison of random and convenience samples of older adults. *International Journal of Behavioral Development*, *26*(4), 345–359. https://doi.org/10.1080/01650250143000247

Irwin, E. G. (2002). The Effects of Open Space on Residential Property Values. *Land Economics*, *78*(4), 465–480. https://doi.org/10.2307/3146847

Jacobe, D. (2013, March 27). *Americans Want More Emphasis on Solar, Wind, Natural Gas*. Gallup.Com. https://news.gallup.com/poll/161519/americans-emphasis-solar-wind-natural-gas.aspx

Jannink, I., Bennen, J. N., Blaauw, J., van Diest, P. J., & Baak, J. P. A. (1995). At convenience and systematic random sampling: Effects on the prognostic value of nuclear area assessments in breast cancer patients. *Breast Cancer Research and Treatment*, *36*(1), 55–60. https://doi.org/10.1007/BF00690185

Jarvis, S. (2021). *The economic costs of NIMBYism: Evidence from renewable energy projects*. 2.

Jensen, C. U., Panduro, T. E., Lundhede, T. H., Nielsen, A. S. E., Dalsgaard, M., & Thorsen, B. J. (2018). The impact of on-shore and off-shore wind turbine farms on property prices. *Energy Policy*, *116*, 50–59. https://doi.org/10.1016/j.enpol.2018.01.046

Jin, J., He, R., Wang, W., & Gong, H. (2018). Valuing cultivated land protection: A contingent valuation and choice experiment study in China. *Land Use Policy, 74*, 214–219. https://doi.org/10.1016/j.landusepol.2017.09.023
Kim, K. J., Lee, H., & Koo, Y. (2020). Research on local acceptance cost of renewable energy in South Korea: A case study of photovoltaic and wind power projects. *Energy Policy, 144*, 111684. https://doi.org/10.1016/j.enpol.2020.111684
Knight, P., Odom, C., Camp, E., Bhandari, D., & Frost, J. (2020). *Solar Siting Opportunities for Rhode Island: An analysis of potentials and costs of rooftop, landfill, gravel pit, brownfield, commercial and industrial ground-mounted and carport solar* [Report prepared for Rhode Island Office of Energy Resources]. Synapse Energy Economics. http://www.energy.ri.gov/documents/renewable/Solar%20Siting%20Opportunities%20for%20Rhode%20Island.pdf
Krohn, S., & Damborg, S. (1999). On public attitudes towards wind power. *Renewable Energy*, *16*(1), 954–960. https://doi.org/10.1016/S0960-1481(98)00339-5
Kuffner, A. (2018, March 16). *Worry over solar sprawl spreads across Rhode Island*. Providencejournal.Com. https://www.providencejournal.com/news/20180316/worry-over-solar-sprawl-spreads-across-rhode-island
Kuminoff, N. V., & Pope, J. C. (2014). Do "Capitalization Effects" for Public Goods Reveal the Public's Willingness to Pay? *International Economic Review, 55*(4), 1227–1250. https://doi.org/10.1111/iere.12088
Lang, C. (2018). Assessing the efficiency of local open space provision. *Journal of Public Economics*, *158*, 12–24. https://doi.org/10.1016/j.jpubeco.2017.12.007
Lang, C., & Cavanagh, P. (2018). Incomplete Information and Adverse Impacts of Environmental Cleanup. *Land Economics*, *94*(3), 386–404. https://doi.org/10.3368/le.94.3.386
Lang, C., Gaur, V., Howard, G., & Quainoo, R. (2021). Incorporating Resident Preferences into Policy Recommendations for Utility-Scale Solar Siting in Rhode Island. *Environmental and Natural Resource Economics Working Papers*. https://digitalcommons.uri.edu/enre_working_papers/4
Lang, C., Opaluch, J. J., & Sfinarolakis, G. (2014). The windy city: Property value impacts of wind turbines in an urban setting. *Energy Economics, 44,* 413–421. https://doi.org/10.1016/j.eneco.2014.05.010
Lehtonen, E., Kuuluvainen, J., Pouta, E., Rekola, M., & Li, C.-Z. (2003). Non-market benefits of forest conservation in southern Finland. *Environmental Science & Policy*, *6*(3), 195–204. https://doi.org/10.1016/S1462-9011(03)00035-2
Loomis, J. B. (1987). Expanding Contingent Value Sample Estimates to Aggregate Benefit Estimates: Current Practices and Proposed Solutions. *Land Economics, 63*(4), 396–402. https://doi.org/10.2307/3146296
Louviere, J. J. (2001). What If Consumer Experiments Impact Variances as well as Means? Response Variability as a Behavioral Phenomenon. *Journal of Consumer Research, 28*(3), 506–511. https://doi.org/10.1086/323739
Louviere, J., Street, D., Carson, R., Ainslie, A., Deshazo, J. R., Cameron, T., Hensher, D., Kohn, R., & Marley, T. (2002). Dissecting the Random Component of

Utility. *Marketing Letters*, *13*(3), 177–193. https://doi.org/10.1023/A:1020258402210
Lovich, J. E., & Ennen, J. R. (2011). Wildlife Conservation and Solar Energy Development in the Desert Southwest, United States. *BioScience*, *61*(12), 982–992. https://doi.org/10.1525/bio.2011.61.12.8
Ma, C., Polyakov, M., & Pandit, R. (2016). Capitalisation of residential solar photovoltaic systems in Western Australia. *Australian Journal of Agricultural and Resource Economics*, *60*(3), 366–385. https://doi.org/10.1111/1467-8489.12126
Manski, C. F. (1977). The structure of random utility models. *Theory and Decision*, *8*(3), 229–254. https://doi.org/10.1007/BF00133443
MA-Smart Solar. (n.d.). Retrieved April 17, 2021, from https://masmartsolar.com/
McDonald, J., MacInnis, B., & Krosnick, J. A. (2020). *Surveying American Public Opinion on Climate Change and the Environment*. 53.
McFadden, D. (1974). *Conditional logit analysis of qualitative choice behavior*. https://eml.berkeley.edu/reprints/mcfadden/zarembka.pdf
McMaster, H. S., LeardMann, C. A., Speigle, S., Dillman, D. A., Stander, V., Pflieger, J., Carballo, C., Powell, T., Woodall, K., Sun, E., Bauer, L., Lee, W., Corry, N., Williams, C., Fairbank, J., Murphy, R., Briggs-King, E., Gerrity, E., Lee, R., & for the Millennium Cohort Family Study Team. (2017). An experimental comparison of web-push vs. Paper-only survey procedures for conducting an in-depth health survey of military spouses. *BMC Medical Research Methodology*, *17*(1), 73. https://doi.org/10.1186/s12874-017-0337-1
Millar, M. M., & Dillman, D. A. (2011). Improving Response to Web and Mixed-Mode Surveys. *Public Opinion Quarterly*, *75*(2), 249–269. https://doi.org/10.1093/poq/nfr003
Mitchell, R. C., & Carson, R. T. (2013). *Using Surveys to Value Public Goods: The Contingent Valuation Method*. RFF Press. https://doi.org/10.4324/9781315060569
Muehlenbachs, L., Spiller, E., & Timmins, C. (2015). The Housing Market Impacts of Shale Gas Development. *American Economic Review*, *105*(12), 3633–3659. https://doi.org/10.1257/aer.20140079
Naylor, D. (2019, October 16). *West Greenwich residents air concerns over proposed solar project*. Providencejournal.Com. https://www.providencejournal.com/news/20191016/west-greenwich-residents-air-concerns-over-proposed-solar-project
Oehlmann, M., Glenk, K., Lloyd-Smith, P., & Meyerhoff, J. (2021). Quantifying landscape externalities of renewable energy development: Implications of attribute cut-offs in choice experiments. *Resource and Energy Economics*, 101240. https://doi.org/10.1016/j.reseneeco.2021.101240
Ong, S., Campbell, C., Denholm, P., Margolis, R., & Heath, G. (2013). *Land-Use Requirements for Solar Power Plants in the United States* (NREL/TP-6A20-56290, 1086349). https://doi.org/10.2172/1086349
Özdemir, R. S., St. Louis, K. O., & Topbaş, S. (2011). Public attitudes toward stuttering in Turkey: Probability versus convenience sampling. *Journal of*

Fluency Disorders, *36*(4), 262–267. https://doi.org/10.1016/j.jfludis.2011.01.003
Pearson-Merkowitz, S., & Lang, C. (2020). Smart Growth at the Ballot Box: Understanding Voting on Affordable Housing and Land Management Referendums. *Urban Affairs Review*, *56*(6), 1848–1875. https://doi.org/10.1177/1078087419861430
Perea, A., Smith, C., Davis, M., Sun, X., White, B., Cox, M., Curtin, G., Rumery, S., Holm, A., Goldstein, R., & Baca, J. (2020). *U.S. Solar Market Insight Executive summary*. Wood Mackenzie and Solar Energy Industries Association.
Pew Research Center. (2019, November 25). *U.S. Public Views on Climate and Energy*. https://www.pewresearch.org/science/2019/11/25/u-s-public-views-on-climate-and-energy/
Pew Research Center. (2020). *Two-Thirds of Americans Think Government Should Do More on Climate*. https://www.pewresearch.org/science/2020/06/23/two-thirds-of-americans-think-government-should-do-more-on-climate/
Qi, L., & Zhang, Y. (2017). Effects of solar photovoltaic technology on the environment in China. *Environmental Science and Pollution Research*, *24*(28), 22133–22142. https://doi.org/10.1007/s11356-017-9987-0
Qiu, Y., Wang, Y. D., & Wang, J. (2017). Soak up the sun: Impact of solar energy systems on residential home values in Arizona. *Energy Economics*, *66*, 328–336. https://doi.org/10.1016/j.eneco.2017.07.001
RIPUC. (n.d.). State of Rhode Island Public Utilities Commission. Retrieved April 17, 2021, from http://www.ripuc.ri.gov/eventsactions/docket/5088page.html
Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, *82*(1), 34–55. https://doi.org/10.1086/260169
Russell, A., Firestone, J., Bidwell, D., & Gardner, M. (2020). Place meaning and consistency with offshore wind: An island and coastal tale. *Renewable and Sustainable Energy Reviews*, *132*, 110044. https://doi.org/10.1016/j.rser.2020.110044
Shoyama, K., Managi, S., & Yamagata, Y. (2013). Public preferences for biodiversity conservation and climate-change mitigation: A choice experiment using ecosystem services indicators. *Land Use Policy*, *34*, 282–293. https://doi.org/10.1016/j.landusepol.2013.04.003
Sunak, Y., & Madlener, R. (2016). The impact of wind farm visibility on property values: A spatial difference-in-differences analysis. *Energy Economics*, *55*, 79–91. https://doi.org/10.1016/j.eneco.2015.12.025
Trainor, A. M., McDonald, R. I., & Fargione, J. (2016). Energy Sprawl Is the Largest Driver of Land Use Change in United States. *PLOS ONE*, *11*(9), e0162269. https://doi.org/10.1371/journal.pone.0162269
Tsoutsos, T., Frantzeskaki, N., & Gekas, V. (2005). Environmental impacts from the solar energy technologies. *Energy Policy*, *33*(3), 289–296. https://doi.org/10.1016/S0301-4215(03)00241-6

Turney, D., & Fthenakis, V. (2011). Environmental impacts from the installation and operation of large-scale solar power plants. *Renewable and Sustainable Energy Reviews*, *15*(6), 3261–3270. https://doi.org/10.1016/j.rser.2011.04.023

U.S. Energy Information Administration (EIA). (n.d.). Retrieved June 24, 2020, from https://www.eia.gov/energyexplained/electricity/electricity-in-the-us-generation-capacity-and-sales.php

US EPA. (n.d.). *The Social Cost of Carbon* [Reports and Assessments]. Retrieved July 9, 2020, from /climatechange/social-cost-carbon

von Haefen, R. H., Massey, D. M., & Adamowicz, W. L. (2005). Serial Nonparticipation in Repeated Discrete Choice Models. *American Journal of Agricultural Economics*, *87*(4), 1061–1076. https://doi.org/10.1111/j.1467-8276.2005.00794.x

Vyn, R. J. (2018). Property Value Impacts of Wind Turbines and the Influence of Attitudes toward Wind Energy. *Land Economics*, *94*(4), 496–516. https://doi.org/10.3368/le.94.4.496

Walsh, R. G., Loomis, J. B., & Gillman, R. A. (1984). Valuing Option, Existence, and Bequest Demands for Wilderness. *Land Economics*, *60*(1), 14–29. https://doi.org/10.2307/3146089

Wee, S. (2016). The effect of residential solar photovoltaic systems on home value: A case study of Hawai‘i. *Renewable Energy*, *91*, 282–292. https://doi.org/10.1016/j.renene.2016.01.059

Wester-Herber, M. (2004). Underlying concerns in land-use conflicts—The role of place-identity in risk perception. *Environmental Science & Policy*, *7*(2), 109–116. https://doi.org/10.1016/j.envsci.2003.12.001

Whitehead, J. C. (1991). Environmental Interest Group Behavior and Self-Selection Bias in Contingent Valuation Mail Surveys. *Growth and Change*, *22*(1), 10–20. https://doi.org/10.1111/j.1468-2257.1991.tb00538.x

Yang, H.-J., Lim, S.-Y., & Yoo, S.-H. (2017). The Environmental Costs of Photovoltaic Power Plants in South Korea: A Choice Experiment Study. *Sustainability*, *9*(10), 1773. https://doi.org/10.3390/su9101773